CONTINUOUS WAVE WIRELESS TELEGRAPHY

BY

W. H. ECCLES, D.Sc.

DEAN OF AND PROFESSOR OF ELECTRICAL ENGINEERING IN THE CITY AND GUILDS OF LONDON TECHNICAL COLLEGE, FINSBURY, LONDON
VICE-PRESIDENT OF THE INSTITUTION OF ELECTRICAL ENGINEERS
VICE-PRESIDENT OF THE PHYSICAL SOCIETY OF LONDON

PART I

ALL RIGHTS RESERVED

LONDON
THE WIRELESS PRESS, LTD.
12 & 13 HENRIETTA STREET, STRAND, W.C. 2
NEW YORK, WIRELESS PRESS INC., 326 BROADWAY.

3733.

For many parts of Nature can neither be invented with sufficient subtilty, nor demonstrated with sufficient perspicuity, nor accommodated unto use with sufficient dexterity, without the aid and intervening of the mathematics.—FRANCIS BACON: "Advancement of Learning."

PREFACE

In this book the author presents in simple form a reasoned :position of the science of Continuous Wave Wireless Telegraphy. In the volume now issued, which constitutes the First Part of ıe treatise, the fundamental principles of electro-magnetism are ›t forth in a manner immediately applicable to the study of /ireless Telegraphy, and the mathematical portions constitute sine and cosine calculus which is especially suitable for ontinuous Wave problems. The treatment of the ionic tube is milarly fundamental, and in the present volume is mainly evoted to the discussion of physical properties, and not of pplications.

The author's thanks are due to Mr. J. Lister and to Dr. J. H. 'incent for reading and criticising portions of the volume, nd to Mr. P. R. Coursey for his generous and efficient help in reparing the work for the press.

W. H. E.

January, 1921.

CONTENTS

CHAPTER IV

CHAPTER V

CONTINUOUS WAVE WIRELESS TELEGRAPHY

CHAPTER I

HISTORICAL SUMMARY

PRODUCTION OF ELECTRIC WAVES

1. OF the many electric methods of transmitting signals from one place to another without the aid of connecting line wires, the method known variously as Wireless or Hertzian or Radio-Telegraphy and Telephony is the only one that has proved of practical utility over long ranges up to this date. This is universally known *par excellence* as " Wireless " Telegraphy and Telephony. This method of signalling is based on the possibility of creating powerful electric waves capable of travelling over the

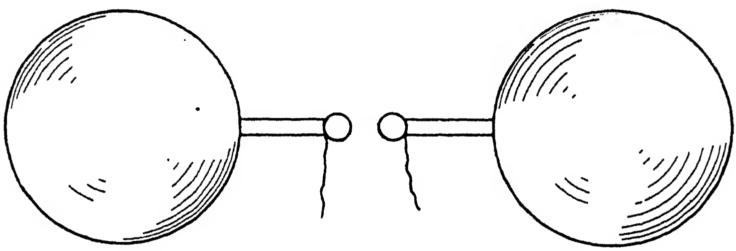


FIG. 1.—Hertz Oscillator.

earth's surface to great distances, and of being detected even when very enfeebled. The feasibility of signalling by aid of electric waves came prominently into view when H. Hertz, in 1887, showed that Clerk Maxwell's visions of propagation of electromagnetic disturbances could be realised. It had long been known that alternating currents of great rapidity could be produced by the abrupt discharge of a charged condenser through a coil of wire; Oliver Lodge had shown that these currents could be used

to produce stationary electric waves on wires; Hertz showed that by the use of a suitable form of charged condenser electric waves could be produced which would spread in all directions from the source with the speed of light. This electric wave producer is now usually referred to as a Hertzian oscillator, and is sketched in Fig. 1. For detecting the waves at a distance Hertz used a simple instrument which he called a "resonator," which consisted merely of a wire ring interrupted by a very small spark gap, as shown in Fig. 2. The waves revealed themselves by producing in this resonator cumulative alternating voltages that produced sparks in the gap.

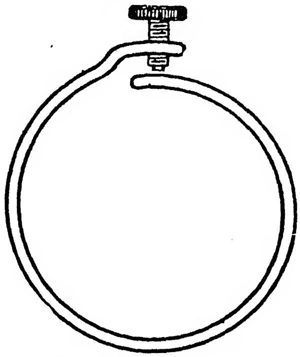


FIG. 2.—Hertz Resonator.

DETECTION OF ELECTRIC WAVES

2. The apparatus of Hertz was on too small a scale and the resonator was too insensitive to permit of signalling over more than a few yards. This latter defect was soon removed by other workers, notably by Lodge and Branly. In 1890 E. Branly showed that imperfect electrical contacts between the particles in a mass of metal filings are greatly improved in electrical conductance when a circuit including the filings becomes the seat of such electrical oscillations as occurred in the resonator of Hertz. Shortly before this, Oliver Lodge, investigating the single contact formed between a metal needle and a plate of metal, found likewise that improvements in the conductance of the contact were produced by electric waves. From his observations he concluded that the forces of cohesion known to exist between very closely approximated portions of matter were stimulated by the electrical agencies, and therefore he called every such device a "coherer."

The increased conductance so produced is easily perceptible by ordinary methods of testing, even when the alternating currents are a thousand times smaller than those needed to produce the visible spark used in the Hertz resonator. Obviously these new modes of detection brought wireless signalling much nearer to practical accomplishment.

3. Lodge and other experimenters, with these and similar detectors, found that it was advantageous to use, for catching the waves, a piece of apparatus resembling the oscillator in shape, called an open or winged resonator, instead of the original circular resonator of Hertz, the detector being placed in the position of the spark gap in Hertz's oscillator. With such apparatus, Lodge, in 1894, transmitted signals a distance of 60 yards, using as one of his accessory pieces of apparatus a bell-trembler mechanism, automatically operated, for setting the coherer afresh after the registration of each signal.

Natural Electric Waves

4. Till 1895, the coherer—under which term must be included the multiple contact arrangement of Branly as well as the single contact device of Lodge—was used merely in connection with wave-making apparatus of laboratory size, that is, of magnitude and power of the same order as Hertz's original oscillator. About this time, however, Popoff applied the coherer and certain of Lodge's accessories to the study of atmospheric electricity, and instead of using a circuit of the size of a Hertz resonator to pick up electric waves and act on the coherer, he used a tall straight lightning conductor. This is shown in Fig. 3, where also the galvanometer G and the voltaic cell E are shown connected in such a manner that when the conductance of the coherer D improves, the previously small deflection of the galvanometer becomes a large one. Popoff's receiving resonator, namely, the long vertical wire, is practically the antenna or aerial of to-day. He found that the coherer was frequently actuated in thundery weather, and not so often during weather free from electrical disturbances. Undoubtedly the coherer connected in series with the lightning rod in these experiments was often operated by true electric waves coming from a distant lightning stroke. Popoff, in describing his experiments on atmospheric electricity with this antenna and coherer, remarked that his disposition of apparatus would be appropriate for receiving preconcerted signals from a distance whenever means were devised for creating and despatching electric waves of suitable type and power. This desideratum was supplied in the years 1895—1896 by G. Marconi, and then Wireless Telegraphy on an engineering scale made its

beginning. It has been stated that at about the same date Henry Jackson, an officer in the British Navy, also succeeded in

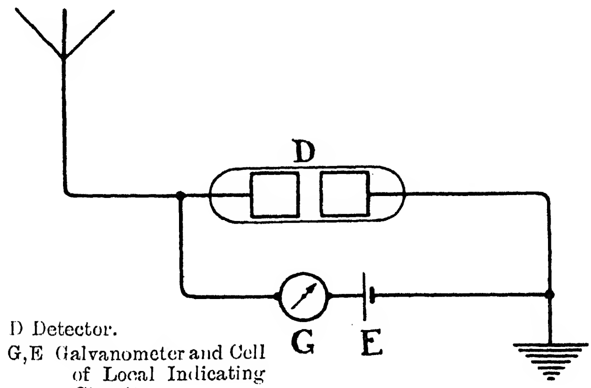


FIG. 3.—Primitive arrangement of Antenna and Detector.

producing electric waves on a practical scale and used them in naval signalling.

PRACTICAL TELEGRAPHY

5. Marconi modified the oscillator of Hertz in the same way that Popoff had altered the resonator; he lengthened one side or wing of the oscillator and disposed it vertically, suppressed the other wing, and connected one spark ball as directly as possible to the earth. The state of affairs at this epoch is expressed in Fig. 4, which gives the essence of Marconi's apparatus diagrammatically.

On the left is the modified Hertzian oscillator with spark gap; on the right is the modified open resonator with the sensitive detector. When the sending antenna A_1 is charged to such a high electrical potential above the earth's that a spark jumps across the gap, electric waves are generated as in Hertz's experiments, and spread outwards from A_1 with the velocity of light. These waves, falling on the antenna A_2, affect the detector, increase its conductance, and enable the arrival of the group of waves to be chronicled at leisure by ordinary slow-moving electrical instruments. It may be mentioned here that Marconi, besides making the large-scale application of the Hertz oscillator,

ery greatly improved every detail of the receiving apparatus id raised the filings coherer to the status of a comparatively ustworthy instrument. When all this had been achieved we ay say that the scientific discovery of the preceding ten years

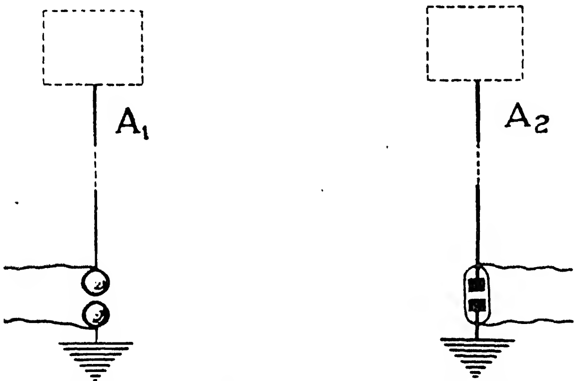


FIG. 4.—Primitive Sending and Receiving Apparatus.

ad now yielded a practical method of communication giving bvious promise of utility in the workaday world—the inquiries f the philosophers had led up to the performance of the engineer.

AN ANALOGY

6. A fuller understanding of some of the electrical processes ıay be reached by considering an analogy. Imagine a surface f still water across which signals are to be transmitted by means f waves on the water. For producing waves on the surface here are many methods available—for example, a stone might e thrown into the water or a stick might be used in various ways o disturb the water surface. For detecting the waves after they ave travelled to the receiving station the small movements of a oating object might be observed by aid of a magnifying appaatus. The splash caused by a stone or by a jerk of the stick ould produce a solitary wave or a short train of waves which ould spread in all directions over the water surface in crests nd troughs forming concentric circles. A longer train of waves an be produced by a sustained regular to-and-fro motion of the tick; but, leaving this possibility for a moment, the state of rireless telegraphy at the epoch under consideration could

probably be best represented by supposing that the stick was used merely to strike the surface of the water once.

Tuning. Wavelength. Damping

7. In the year 1897 Lodge filed a patent specification in which the advantages of giving to sender and receiver the same definite natural frequency were explained, and in which means of carrying out modes of effecting the adjustment were described. The analogy of the waves on the still water surface will at this stage be more perfect if we suppose the wavemaking and receiving apparatus to have definite natural frequencies. Suppose then that the sending apparatus is a moored buoy floating freely on the water, and that the receiving apparatus is a similar buoy at a distance. Such floating buoys have quite definite frequencies of rise and fall in the water. If the sending buoy be pushed deeper into the water and suddenly released it will rise and fall periodically with gradually diminishing excursions, and will create a series of humps and hollows in the water round it. These will travel outwards over the water surface as concentric rings of gradually diminishing height.

The time of a complete to-and-fro motion of the buoy is called the period of its oscillation. In the first period of the motion of the buoy a wave consisting of a trough followed by a crest will start diverging across the pond. In the second period another trough and crest will be despatched, and so on for successive oscillations of the buoy, as shown in Fig. 5. The distance between the first trough and the second, or between the first crest and the second, or between the second trough and the third, and so on, will be the distance travelled by the waves in the time of one period of vibration of the buoy. This distance is constant at any particular frequency, and is called the wavelength corresponding to the period of oscillation. The energy of the motion given to the water is, of course, drawn from the original stock of potential energy given to the buoy in pushing it into its initial depressed position, and it is clear that the despatch of waves, or, in other words, the act of radiating, gradually robs the buoy of its stock of energy. This loss by radiation causes the second oscillation to be smaller than the first, the third than the second; and, therefore, the first wave is larger than the second, the

second larger than the third, and so on. Such series of decaying oscillations and trains of waves are commonly said to be "damped."

It is supposed in this analogy that the velocity of propagation of waves over the surface is the same whether the waves are of large or small amplitude. In that case the train of damped waves will travel across the water surface without alteration in the distance between any definite pair of crests or troughs. Each wave will, however, diminish in amplitude as it travels; first, because of the divergence in all directions which, so to speak, distributes the energy over steadily increasing areas, and, second, because of the frictional losses arising from the

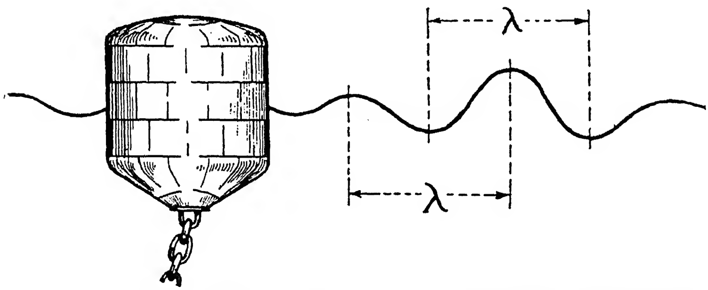


FIG. 5.—Damped train of waves made by buoy depressed and released. λ is the wavelength.

viscosity of the water. Thus the train of damped waves will reach the receiving point much weakened, but the crests and troughs will pass that point in just the same time as elapsed in their creation.

8. A floating mote at the point will, therefore, undergo a series of up-and-down movements that will be a reproduction in miniature of the motion of the buoy. If we suppose the floating mote to be watched through a microscope we have a state of things representing the condition of Wireless Telegraph practice in the year 1898. But after the publication of Lodge's patent the condition of affairs would be better represented by supposing that, instead of a light mote, we have in our analogy a small floating buoy possessing a frequency of its own. This makes the problem of the motion of the receiver more complicated, but at this stage we may assume that the greatest aggregate motion of the receiving buoy will be attained when its natural frequency is the same

as that of the sending buoy, a sufficient reason being that the successive waves of the group emitted by the sender will, in this event, act on the receiver in correct time relation for each to add to, rather than subtract from, the effect of its predecessor. Along with this advantage of greater aggregate effect on the receiver, there is another which has been associated throughout the literature of the subject with the words "syntony" and "tuning." For it is evident that other waves, on reaching the receiving buoy, if they have a different frequency from it, will not possess the cumulative property expressed above, and will, therefore, produce little effect on the resonator. That is to say, the tuned receiver is selective, or, in other words, responds readily to waves of its own frequency and only slightly to waves of very different frequency.

Damped Electric Waves

9. To translate the above analogy into the electrical terms appropriate to Wireless Telegraphy with damped waves we must replace the floating buoy at the sending end by a Hertzian oscillator, or its practical equivalent the Lodge antenna, or, less precisely, the Marconi plain aerial. Such an oscillator, when charged so that a high voltage arises across the spark gap, discharges suddenly when the gap breaks down, and the current, finding the gap a good conductor when filled with the spark, attains its maximum value by the time discharge is finished. The current does not stop, but continues and charges the antenna in the reverse sense, then stops, reverses, and goes through the same programme in the next half period of the oscillation which is thus set up. In fact, by this discharge a damped train of oscillations is created whose energy content depends on the initial voltage and on the capacity of the antenna, and whose frequency depends also on other electrical qualities. Part of the energy is wasted in local resistance, both in the antenna and in the ground, but some is detached in the form of electric waves just as were the waves in the water analogy. There is the difference, however, that the electric waves diverge in three dimensions, and not merely in two, which makes pictorial representation of the state of affairs very much more difficult.

The train of damped waves spreads with the speed of light in all directions, but with most intensity in the horizontal direction.

The intensity of the beam falls off in every direction according to the inverse square law that governs all radiation in three dimensions, or perhaps rather more quickly than this on account of the absorption in the air and in the materials of the earth's crust. Thus the waves reach the receiving antenna much attenuated, but with practically the same damping; in other words, a train contains the same number of waves as at the start, but each wave is very much feebler.

10. Following the analogy still further, and assuming that the receiving antenna is, like the buoy, tuned to the incoming waves, these then excite by resonance a small oscillatory current in the aerial. To describe the mode of detection we must depart from the analogy, and, indeed, to describe the process fully is the purpose of a later section of this book; but here we may say that the energy of this oscillatory motion of electricity in the receiving antenna is drained away to the detector by appropriate electrical connections probably almost as fast as it is collected by the antenna, and a fraction of it is converted by the properties of the detector into a quick pulse of unidirectional current, which passes round the telephone receiver circuit, causes a movement of the telephone diaphragm, and so affects the ear drum of the operator.

Musical Signals

11. Since each spark produces a train of waves, and each train reproduces itself in the receiving antenna as a damped train of oscillations which are converted more or less imperfectly by the detector into a pulse of telephone current, it is clear that a sequence of sparks at, for example, the rate of 256 per second will produce 256 pulses of unidirectional current through the telephones and give to the ear an impression of a note of about the same pitch as the middle C of the pianoforte. Now the ear is a sensory organ of energy type, and in the case of sustained tones the intensity of the sensation of sound is, therefore, within limits, proportional to the rate of supply of energy, that is, to the power. From this point of view it is an advantage to use a quickly sparking rather than a slowly sparking sender up to that frequency at which the sensitiveness of the ear begins to fall away; but, in any case, other circumstances being the same, the power spent

at the sending station and the power received by the operator's ear at the receiving station may be taken as proportional in the first approximation. This points the importance in our subject of the physical quantity called "power."

12. The receiving antenna may be regarded as an electrical ear trumpet reaching many feet into the air and catching some of the energy travelling in the form of electric waves through the space near it. Evidently for successful signalling by methods using the telephone receiver the rate of collecting energy must be at least sufficiently great to produce a sound that is audible. Estimates of the minimum audible power have been made by various observers and have agreed fairly well; some experiments of Rayleigh showed that a tuning fork of 256 frequency was barely audible to the untrained ear at a distance of 27·4 m over level ground when it was radiating at the rate of 40 ergs per second. If the auditory meatus be taken as 0·5 sq. cm in cross-section these figures show that the ear can perceive as continuous sound a stream of energy amounting to $0{\cdot}44 \times 10^{-6}$ erg, which, by the way, happens to be of the same order as the power just perceptible to the eye. If the inverse square law be again applied to these figures, we find that a sound just audible to the unassisted ear could be produced at a kilometre by a tuning fork radiating the sound energy at the rate of $5{\cdot}6 \times 10^{-3}$ watt, or at a distance of 100 km by a fork radiating at the rate of 56 watts. With an efficient ear trumpet of aperture 1 sq. m in area the necessary power in the last case would be only 56×10^{-4} watt, and, in fact, if the earth were flat, the power of 5·6 watts would be just audible at 3,300 km. This serves to illustrate the possibilities of Wireless Telegraphy.

Comparison with Acoustic Telegraphy

13. It may be asked why signalling by sound is not employed over great distances. Why is the aid of electricity invoked at all? The reasons are numerous and emphatic, and some of them are worth noting down. First, sound waves are easily reflected and refracted by banks of mist and by strata of cold and warm air; second, they are bent out of their course by the action of the wind; third, their energy is rapidly dissipated by the viscosity of the air; fourth, the sound waves, especially when short, are easily scattered by solid objects; fifth, miscellaneous noises are

everywhere rather abundant; and, sixth, even if all the above difficulties were non-existent, the waves of sound if despatched might not succeed in bending over high mountain ranges or round the protuberance of the globe. As a matter of fact it is not practical, and it would not be pleasant, to make waves of atmospheric compression and rarefaction intense enough to overcome these difficulties.

14. In contrast to all these disadvantages of sound waves we find for electric waves that, first, they are almost entirely unaffected by the variations in humidity or temperature of the atmosphere through which they pass; second, air currents have negligible influence; third, absorption of energy by the air is not nearly so rapid in the electrical case as in the case of sound waves; fourth, electric waves may be made thousands of times as long as sound waves and scattering thus reduced; sixth, the greater wavelength possible for electric waves, or perhaps some purely electrical property of the earth or the atmosphere, enables them to diffract easily over mountains and round the earth. The fifth item of the enumeration of § 13 is one in which it is doubtful whether the electric waves do or do not have any advantage over signalling by sound. Natural electric waves formed most probably by lightning strokes in the atmosphere produce noises known as strays or statics in the telephone receivers of a wireless receiving station, and these are very troublesome every day in tropical stations, and quite frequently in stations in temperate latitudes.

High-Power Apparatus

15. In 1897 Lodge had clearly stated the scientific reasons for the use of inductance coils in tuning antennæ; Marconi, during the years between 1896 and 1900, had been feeling his way by ial trials towards the best practical sizes for every part of his apparatus and had been seeking the wavelengths best for propagation over the earth's surface. This engineering experience and the scientific principles led Marconi, about the year 1900, to introduce into Wireless Telegraphy the conception of a reservoir circuit, which is, in a sense, an equivalent of the capacity areas advocated by Lodge. The reservoir circuit became emphatically necessary about the date mentioned, because, in the ambition to achieve greater ranges, the shocking

coil and the voltaic battery of the Hertz assemblage were being replaced by alternators and transformers of commercial type and rating. This reservoir circuit is merely an intermediate circuit of large electrical capacitance containing the spark gap and practically non-radiative, which takes the energy from the alternator, modifies its voltage and other characteristics, and passes it on in convenient form to the antenna. The reservoir circuit is still in use (1918) in most spark methods.

Energy Considerations

16. The practical result of the whole of the processes occurring from the instant of the discharge and the despatch of the wave train until the instant when the telephone diaphragm at the receiving end is moved, is that the operator situated hundreds of miles away hears each spark, so to speak, within a small fraction of a second of its occurrence. He, therefore, hears the rattle or the musical note, as the case may be, formed by a series of sparks, just as if he were standing within earshot of it. Being given these facts, the task of making intelligible signals by a code employing dots and dashes is an easy one—the sending operator merely causes his apparatus to emit short or long series of sparks, and the receiving operator hears them and records them as dots and dashes respectively. A dot may last one-twelfth of a second, and a dash one-quarter of a second. If, for example, the spark rate is 256 per second, a dot contains about 20 sparks and a dash about 60 sparks, and the receiving operator hears more or less musical notes of the pitch of the middle C of the pianoforte of short and long duration. To enable the sending operator to control the emission of waves he is provided with a Morse key similar to that made familiar by line telegraphy, which makes and breaks at will some convenient link in the chain of the electrical circuits constituting the sending plant.

17. The oscillatory current created in the receiving antenna by the waves represents a quantity of energy absorbed from the waves, and with this stock of energy the signal-recording apparatus has to be operated. Some of it is inevitably wasted—losses occur as joulean heat in the wires and in the earth near the foot of the antenna, and other losses may or may not arise from the fact that the antenna when oscillating radiates back into space some of the energy it has gathered.

To study further the energy quantities it is convenient to start at the telephone receiver. Experiments have shown that the efficiency of this instrument varies enormously with the pitch of the vibrations impressed on the diaphragm, at any rate when loud sounds are being produced. This is an effect quite apart from the physiological fact that the ear has different sensibilities at different pitches. Unfortunately the telephone appears to have very low efficiency at all pitches for the faint sounds used in wireless signalling. We may put it roughly as between one-fifth of 1 per cent. and 3 per cent. Now the current that passes into the telephone receiver comes from the oscillation detector; the detector has converted the antenna's oscillations into the direct current form suitable for the telephone receiver. The efficiency of this conversion in a good detector seems to be about 3 per cent. The question is now: What fraction of the energy collected by the aerial is delivered *via* the detector to the telephone circuit and ultimately to the ear? Published results on this important point are rare. The author, in some of his own experiments, arrived at the following figures. When energy was supplied to the antenna at the rate of $0{\cdot}33 \times 10^{-8}$ watt, *i.e.*, 33,000 millionths of an erg per second, from a discharge apparatus producing a musical note of the rather low frequency of 50 vibrations per second there was a very faint sound heard in the telephone. Assuming the energy coming as sound waves from the telephone diaphragm to the ear to be about 5×10^{-6} erg per second, we find that the efficiency of the whole arrangement of antenna, detector and telephone is $5 \times 10^{-6} \div 33{,}000 \times 10^{-6}$, which equals 0·015 of 1 per cent.

Effects of Rate of Sparking

18. In the case of an oscillator whose spark gap breaks down always at the same charging voltage we have seen in § 11 that more power will be conveyed from sender to receiver when the sparks are rapid than when they are slow. For the sake of the receiving operator, and also to satisfy certain mechanical predilections of the telephone diaphragm, it is better that the sparks should be regular and of audible frequency, or, in other words, that they should make in the telephone a musical sound of constant medium pitch. Suppose that the frequency of the

electrical oscillations is a million per second, which corresponds to a wavelength of 300 m, and that the sparks occur at the rate of 256 per second; suppose further that the rate of damping is such that the energy left in the train of waves after the first 20 have been counted is negligible; then since 20 times the period of one oscillation is 20 millionths of a second and the time between the sparks is nearly 4,000 millionths of a second, we see that 20 portions of time are occupied by oscillations and 3,980 portions unoccupied in the whole time between any two sparks. That is to say, the idle interval is nearly 200 times as long as the active interval. This is from many points of view the main defect of the ordinary spark method of wireless telegraphy. The defect has been combated somewhat by the introduction of the so-called quenched spark, which is essentially a method for compressing into the suggested interval of one two-hundred-and-fifty-sixth part of a second a long series of very rapid sparks occurring at a rate of possibly 10,000 per second.

Quenched Sparks

19. The additional improvement that may be brought by the introduction of what is called the quenched spark will be appreciated from a consideration of the diagrams of Figs. 6 and 7, which exhibit the contrast between it and the ordinary spark method. The sine outline represents the voltage of an alternator making 500 alternations per second. In the ordinary spark there are two sparks occurring in the alternator's period, each setting its circuit into oscillations having a frequency of 200,000 per second, which are prolonged through a time of one four-thousandth of a second. In the quenched spark diagram eight or nine discharges are shown in each half-period, and these occur at intervals of one ten-thousandth of a second and give short trains of oscillations. In both figures the diagrams relate not to the oscillations of the antenna, but rather to those of the reservoir circuit. It will be seen that while in Fig. 6 the idle time may be as much as 90 per cent. of the whole, in Fig. 7 it may be only 40 per cent. Thus if each train is of the same energy content, and if the relative losses in all the auxiliary circuits are the same in the two systems, there is more energy emitted by the quenched spark than by the ordinary spark in the same time. If, on the other hand, it is required to emit the same power, the quenched

spark accomplishes the task with lower voltages than the ordinary spark. Certain difficulties inherent in quenched spark working make the latter condition the more usual one.

The trains are emitted in gushes at the rate of 256 gushes per second, and these produce in the receiving apparatus a note of that pitch; but the actual discharges within the gushes are too rapid to be easily audible, though they make themselves manifest as a hissing sound and give a breathy quality to the audible note.

Even with quenched sparks making 10,000 discharges per

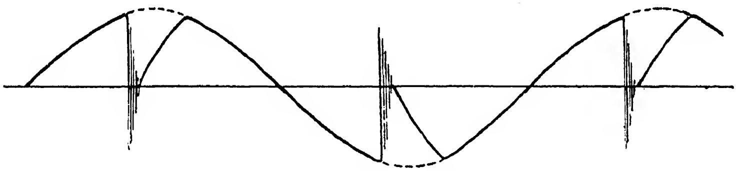


FIG. 6.—Diagrammatic representations of voltage of antenna when excited to give 1 spark per reversal of the alternating supply voltage.

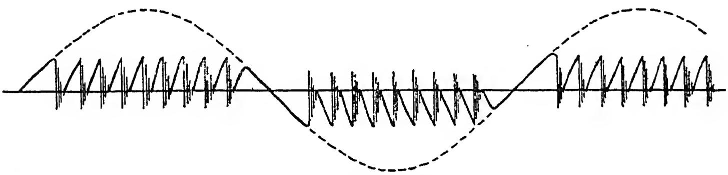


FIG. 7.—Diagrammatic representations of voltage of antenna when excited to give 9 sparks per reversal of the alternating supply voltage.

second the idle time is, in the above example, as much as 40 per cent. of the whole. Moreover, each train of waves is damped and therefore a part of the active time is rather feebly utilised. Clearly, additional improvement may still be hoped for in the further diminution of idle time and in the equalising of the oscillations. In small plants certain inventors have obtained a degree of success by increasing the rapidity of the discharge till the beginning of each train trespasses on the space of time belonging to the preceding train. But evidently the ideal solution of the whole problem would consist in filling up the whole time with high-frequency sine form alternating current. Then, when the signalling key is depressed, the antenna is excited by sustained or continuous oscillations, as high-frequency alternating current is called, and continuous waves are radiated.

Continuous Wave Telegraphy and Telephony

20. New achievements in Wireless Communication now follow inevitably. In the first place the electrical limitations which slow sparks impose upon the rate of making dots and dashes are removed, and high-speed telegraphy employing automatic transmission becomes feasible, even to the extent of hundreds of words per minute. In the second place Wireless Telephony becomes easy and excellent in performance as compared with what could be accomplished by means of the rapid spark methods formerly much experimented with; for the high-frequency alternating current in the antenna is as easily modulated by a microphone as is the direct current employed in ordinary line telephone systems.

Continuous wave radiotelegraphy has been predicted and pursued since the earliest days of wireless telegraphy. But even after the means of generating powerful continuous waves had been found in the Poulsen arc and the high-frequency alternator, the inconveniences and imperfections in the methods of reception at first available prevented the full realisation of the advantages of continuous wave telegraphy and delayed the bestowal of the preference over spark telegraphy that it merits. Since 1912, however, the marvellous potentialities of the three-electrode thermionic vacuum tube have been explored, unfolded, and turned to account in unforeseen and subtle ways, with the result that the new methods far transcend the old in elegance and power and promise, and have enabled electric waves to bear the human voice across the Atlantic Ocean and to carry dots and dashes across three continents from Wales to Australia.

Chapter II

ELECTROSTATICS AND ELECTRODYNAMICS

1. In any account of the principles of Wireless Telegraphy the properties of the pieces of apparatus known as condensers and inductance coils, and the phenomena of the propagation of electric waves, must be referred to so frequently that it seems desirable to set forth at the very beginning a plain statement of the fundamental theorems of electrostatics and electrodynamics in the form most useful for the special requirements of our main subject. This statement will proceed upon the assumption that the reader is already acquainted with experimental electricity, and completeness and rigour will not be aimed at.

ELECTROSTATICS

2. Electricity became evident to our senses mainly because one portion of electricity exerts mechanical force upon another, and in electrostatic theory advantage is taken of this fact in order to define a unit of electricity in terms of which any given portion of electricity can be measured. The reader will remember that the "electrostatic unit charge" is such a quantity that if it could be concentrated at a point and placed 1 cm distant from an equal point charge in a vacuum their mutual repulsion would be 1 dyne. The unit of electricity and a number of derived units can, therefore, be established so soon as a unit of length and a unit of force are settled upon. In the centimetre-gram-second system of units the unit of length is the centimetre, that of force the dyne. The various electrical units are discussed in § 99.

This unit of charge is not the unit used in our ordinary work with electric currents, the coulomb; for the coulomb contains almost exactly 3×10^9 electrostatic units. Neither is it what might be called the natural unit of electricity, the electron; the electrostatic unit contains about $2 \cdot 1 \times 10^9$ electrons. The reader will doubtless know that the electron is the "atom" of negative electricity—the smallest amount of electricity yet recognised in

any electrical event. An electron, if moving at a speed well below that of light, appears to possess inertia less than a thousandth part of that of a hydrogen atom. Positive electricity has not yet been isolated from matter in the same sense that negative has, and since when one electron is removed from any gaseous atom the remainder behaves as if charged positively—and is called a positive ion—it is permissible for many purposes to conceive that a positively charged conductor of molar dimensions is one with too few electrons, while a negatively charged conductor is one with more than the normal number of electrons.

Experiment has shown that if Q units of the kind defined above be placed in vacuum at a distance x cm from Q' units in vacuum, both quantities being concentrated at points, the repulsion or attraction, as the case may be, between the charges will be QQ'/x^2 dynes. This is known as the inverse square law, and from it the mathematician can solve problems such as the distribution of electricity on conductors in an electric field, that is, in the portion of space into which the influence of given electric charges extends.

3. There is no need to explain in this book the difference between conductors and insulators, or to point out that conductors are opaque to the electric effects and that insulators are transparent to them, and are therefore called "dielectrics." But it is well to mention that when insulating matter intervenes between two concentrated charges Q and Q' placed x cm apart the mechanical force between them is reduced. A famous experiment of Faraday's shows that if all the space round the charges were filled with a homogeneous dielectric, such as mineral oil, the mechanical force between the charges would be given by the expression

$$\frac{QQ'}{\kappa x^2}$$

The divisor κ was called by Faraday the specific inductive capacity of the dielectric. It is also called the dielectric constant, the permittivity, the electric inductivity, or simply the inductivity. A useful abbreviation of the first name is S.I.C.

Faraday Lines.

4. The influence of the medium on the law of force caused Faraday to introduce the conception of "lines of force" for

picturing to the mind the electric field. These are lines drawn continuously so that the direction of each point of each line is the direction of the mechanical force that would act upon a concentrated charge if placed there. For definiteness we shall suppose each unit positive charge to have one line starting from it and each negative unit to have one line ending on it. If lines are, as some suppose, real physical entities, there must be at least one line attached to each electron; and since there are $2{\cdot}1 \times 10^9$ electrons in one of our artificial units, this number of physical lines at least must be regarded as represented by one Faraday line of our diagrams and calculations. Since a force cannot have two directions at the same point, we conclude that the physical lines from an electrostatic unit of charge never cross one another, but form in the dielectric a discrete bundle, sometimes called a Faraday bundle or tube. In future, when speaking of a Faraday line, the bundle or tube is meant, and therefore we may be permitted to speak of a fraction of a Faraday line on occasion.

5. The principal statical properties of these lines were pointed out by Faraday himself. They tend to contract lengthwise and to repel one another sideways. Maxwell showed that these tensions and repulsions could have values assigned to them that enabled the forces between electrified bodies to be fully accounted for without the use of the inverse square law. We shall return to this shortly, but may remark now that though the lines are governed by these curious laws of force, it does not follow, and in fact it is not the case, that the dielectric they pass through feels in bulk such forces. And even at the ends of a line of force, where it rests on and apparently pulls at the conductor, we must suppose that the electricity primarily feels the tension, and then, since electricity cannot easily be dragged off a conducting surface, passes the pull to the material body.

Electric Stress and Electric Strain.

6. In order to allow for the effect of the dielectric it is useful to regard the Faraday lines as mapping the electric strain in the dielectric and to suppose this strain to be the same whatever homogeneous dielectric fills the field. The latter hypothesis is necessitated by the definition of Faraday lines, which demands that the Faraday lines shall be unchanged if the charges of elec-

tricity are. Then from mechanical analogy we picture the strain to be due to electric stress. The stress is called electric force by some authors, electric intensity by others ; the symbol F will be used to indicate its value at any point. The strain was called electric induction by Faraday, displacement by Maxwell, polarisation by J. J. Thomson. The strain is represented by the symbol σ, and its value at any point is supposed estimated by the density of the Faraday lines at the point. By density is meant here the number of lines per square centimetre crossing a

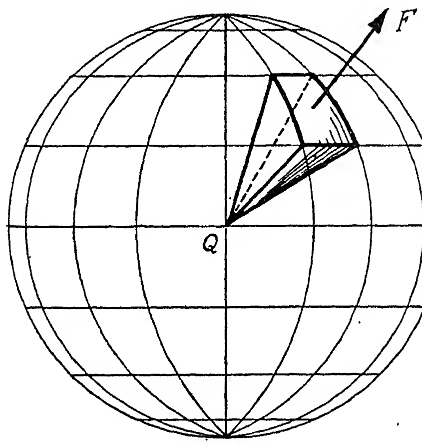


FIG. 8.

small area round the point and perpendicular to the lines. It is quite analogous to flux-density in magnetism (§ 31).

7. The relationship between the stress and strain is best stated by returning to a point charge of electricity of quantity Q esu placed in a dielectric of S.I.C. κ. The Faraday lines, Q in number, spread from it equably in three dimensions (Fig. 8). At any point distant x cm from Q the electric force is

$$F = Q/\kappa x^2$$

according to fundamental experiments. Also by the definition of the electric strain, since the Q lines are spread equably over the area of the sphere radius x, we have

$$\sigma = Q/4\pi x^2$$

for the lines per unit area.

Therefore

$$4\pi\sigma = \kappa F$$

or the electric strain is $\kappa/4\pi$ times the stress.

As an example of these results let a charge of 314 esu be present on a very small conductor in the middle of a large tank full of oil, and let the S.I.C. of the oil be 2. Then the electric force or field or intensity at a distance of 10 cm is

$$F = 314 \div 2 \times 10^2$$
$$= 1{\cdot}57 \text{ esu.}$$

This electric force could be made evident by bringing into the field a "test-charge" of, say, 10 units of electricity carried on a small insulated conductor. If it were placed at the point for which the electric force has been calculated above it would be acted upon by a mechanical push of 15·7 dynes. A test-charge is usually imagined to be 1 unit of electricity and then the mechanical force upon it when it is placed at any point is the measure of the electric intensity or force existing there. The electric force is not perceived, it should be noticed, until the test-charge is brought in; nor is gravity perceived until a piece of matter, such as a stone, is used to show it.

At the point in the oil where the electric force is 1·57 esu the strain σ would be $\sigma = (2 \div 4\pi) \times 1{\cdot}57 = 0{\cdot}25$ Faraday line per square centimetre. It is just as convenient to calculate the strain first; thus

$$\sigma = 314 \div 4\pi \,.\, 10^2 = 1/4 \text{ esu,}$$

as before, and therefore

$$F = (4\pi \div 2)\,(1/4) = \tfrac{1}{2}\pi \text{ esu.}$$

8. In the electrostatic system of units the value of κ for vacuum is taken as 1. It seems possible, or even probable, that when matter is present in the field the positive and negative electrical portions of the molecules, whether the matter is solid, liquid, or gaseous, are pulled slightly apart by the electric force; and this special yielding of the medium enables a given electric force to produce a greater strain in matter than in a vacuum. This displacement of molecular electricity may have led Maxwell to give the name "displacement" to the electrical strain.

Potential.

9. Free positive electricity, if placed in an electric field, finding itself acted upon by the electric force there, tends to move in the

direction of the force just as a stone falls towards the earth under gravitational force. Under gravity the tendency is for the stone to move from where it possesses higher potential energy to where it possesses lower, and the same holds for a free electric charge in an electric field. In any movement of the stone from one place to another, whether natural or forced, the difference of potential energy of the stone is equal to the work done in the movement—the potential energy decreases if gravity on the whole does the work, but increases if the work is done on the whole by the outside agency. A similar statement can be made for the electrical case. Obviously the more intense the forces in the field the greater will be the amount of work done in any stated movement, and thus the difference of potential energy at two places affords a measure of the general condition of the field. For this way of reckoning fields it will clearly be best to have the test-stone or the test-charge always of the same size, for the potential energy loss or gain will evidently be proportional to the magnitude of the body or charge moved; and it is always in fact supposed to be a unit mass or charge. When this is understood the word "energy" is dropped and we speak of potential and of potential difference. Potential difference, often abbreviated to P.D., may be a rise or a fall—when the movement of the test mass or charge is in the natural direction it is a fall, and is often called potential drop. Thus the potential drop between two places in an electric field is work per unit charge done by the field when electricity is moved by the field from one place to the other; or it is the loss of potential energy per unit charge of electricity moving from one place to the other.

10. All the above applies whether the electricity moves as a free charge through a dielectric or as portion of a current through a conductor. For example, if we find by aid of an electrometer or voltmeter that a voltaic cell has a potential of 1·2 volts between its terminals while a steady current flows through a wire joining them, we know that every coulomb of electricity loses 1·2 joules of potential energy in passing from the positive terminal to the negative through the wire. Its potential drop in the wire is 1·2 volts; but its potential rise in passing through the voltaic cell is 1·2 volts, the lift in voltage being accomplished by chemical means and at the expense of chemical energy. Moreover, if a quantity of free positive electricity, carried on an insulated

conductor, were started from very near the positive terminal and moved very near to the negative, the work done per coulomb would be 1·2 joules also.

But the volt is not the electrostatic unit of potential. In this system of units the unit of work is the erg, not the joule, which is 10^7 ergs; and the unit of charge is so small that there are 3×10^9 electrostatic units in one coulomb. Therefore, the electrostatic unit of potential is $3 \times 10^9 \div 10^7$ volts, that is,

$$1 \text{ esu of potential} = 300 \text{ volts}.$$

It may be remarked here that difference of potential between two points on or in a conductor is evidence that there is an

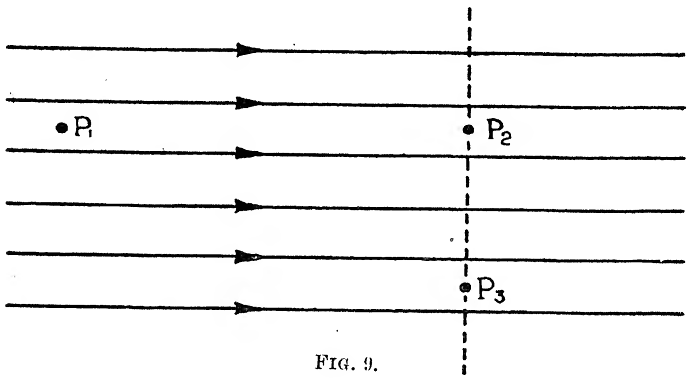


FIG. 9.

electric current running in the conductor, or that there is a source associated with it and creating the potential difference. If there are no sources the current, which is due to the tendency already spoken of for electricity to move so as to diminish its potential, endures only until all electricity in the conductor is at the same level of potential. The surface of any conductor is therefore an equipotential surface after transient currents have subsided.

Potentials in some Typical Fields.

11. The electric fields with which we have to deal later often fall into one of three types, namely, uniform fields, cylindrical fields, and spherical fields. In a uniform field the Faraday lines are all parallel and equidistant (Fig. 9); such a field occurs between two large parallel plates oppositely charged and held near together. A cylindrical field arises round a very long charged isolated rod or wire (Fig. 10), and the Faraday tubes

project perpendicularly from the wire, like the fibres from a round brush; this is called two dimensional divergence. A spherical field appears when a charged sphere or a charged point is placed at a great distance from other matter (Fig. 11); the Faraday lines spread uniformly in three dimensions.

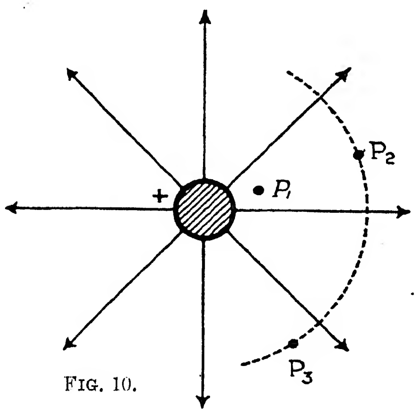


FIG. 10.

12. In a uniform field if the Faraday line density or strain is σ the force is $4\pi\sigma/\kappa$, and the work done by the field in moving a unit charge parallel to the field from P_1 to P_2 is

$$V_{12} = Fx = 4\pi\sigma x/\kappa$$

where x is the distance P_1P_2. This is the potential drop.

On moving the test-charge from P_2 to P_3, perpendicularly to the Faraday lines, there is no work done because there is no component force in that direction. Therefore P_3 is at the same potential as P_2 and $V_{13} = V_{12}$. The surface through P_1 perpendicular to the Faraday lines is an (imaginary) equipotential surface. It should be noticed that if the test-charge were moved from P_2 to P_3 in an erratic path so that sometimes work was being done by the field and sometimes work being done against it, the whole work must be zero by the principle of the conservation of energy.

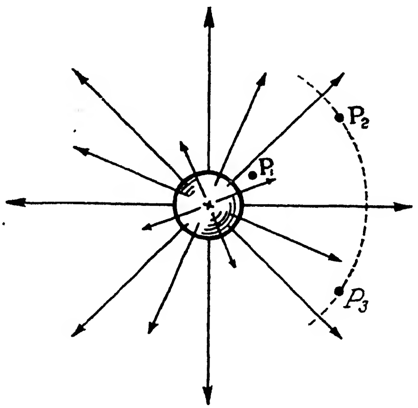


FIG. 11.

13. In a cylindrical field let the charge per centimetre of wire be q_1, the distance of P_1 from the axis be x_1, and the distance of P_2 be x_2. Concentric cylindrical surfaces through P_1 and P_2

would each cut q_1 lines per unit of axial length. Therefore the Faraday line density at P_1 is

$$\sigma_1 = q_1/2\pi x_1$$

The electric force here is

$$F_1 = (4\pi/\kappa)\sigma_1 = 2q_1/\kappa x_1.$$

At any other point, distance x,

$$F = 2q_1/\kappa x$$

and it therefore varies inversely as x.

In order to calculate the work done by the field in pushing the test-charge to P_2 we must integrate, and so obtain for the potential drop

$$V_{12} = \int_{x_1}^{x_2} F dx = \frac{2q_1}{\kappa}\int_{x_1}^{x_2}\frac{dx}{x}$$

$$= \frac{2q_1}{\kappa}\Big[\log_\epsilon x\Big]_{x_1}^{x_2}$$

$$= \frac{2q_1}{\kappa}\log_\epsilon\frac{x_2}{x_1}.$$

This result may be called the potential drop due to two dimensional divergence.

14. In a spherical field the lines spread equably in every direction so that if Q is the total charge on the sphere or at the point we have for the line density at any point distant x the Faraday line density

$$\sigma = Q/4\pi x^2.$$

The force is therefore

$$F = Q/\kappa x^2$$

as we have seen already. The potential drop from P_1 to P_2 is

$$V_{12} = \int_{x_1}^{x_2} F dx$$

$$= \frac{Q}{\kappa}\Big[-\frac{1}{x}\Big]_{x_1}^{x_2} = \frac{Q}{\kappa}\left(\frac{1}{x_1} - \frac{1}{x_2}\right).$$

In both the last cases the points P_2, P_3 are at the same potential because P_3 can be reached by moving always perpendicularly to

the Faraday lines. As before, by the principle of the conservation of energy, we see that whatever the path from P_1 to P_2 the value of V_{12} is as above.

CONDENSERS.

15. If we imagine two conductors of any shape placed far apart in empty space, one being charged with Q units of positive electricity and the other with Q units of negative electricity, the Faraday lines attached to each would spread in all directions

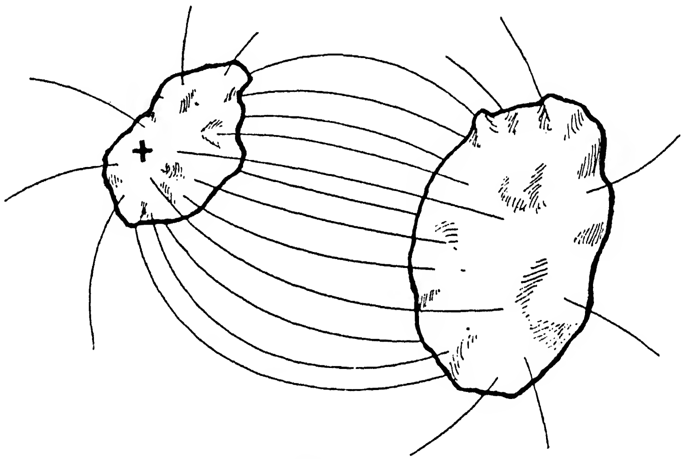


FIG. 12.

almost impartially, though there would be a little bias for some of the lines that start on the positive electricity to bend toward and end upon the negative electricity, and reciprocally. But if the conductors were brought gradually nearer, more and more of the available lines would join up to bridge the space between the conductors, especially on the nearer surfaces, as indicated in Fig. 12. The lines become, so to speak, condensed in the region between the conductors. The conflict between the tension along the lines, which tends to shorten them, and the repulsion between them which tends to lengthen them, is here very evident. Such a pair of conductors brought near together in order to derive certain advantages of this phenomenon constitute a condenser. A condenser is said to be charged when equal quantities of positive and negative electricity are given to the opposite

sides of it. The fact that Faraday lines stretch from one to the other shows that the conductors are at different potentials and, therefore, when the conductors are connected together by a wire a current flows and the charges neutralise each other.

Capacitance.

16. The potential difference between the sides of the condenser is estimated ideally by finding the work done on a mobile unit test-charge when it is pushed by the field from one conductor to the other. The principle of the conservation of energy ensures that this will be the same whatever the starting or finishing point or the path of the test-charge. We shall suppose the potential difference V to exist when each charge is of magnitude Q. Now let every portion of the electricity on the conductors be added to till each is doubled ; the Faraday lines will of course then be twice as numerous. We shall assume that the electricity and the lines will be at rest in positions geometrically similar to the old positions—a proposition falling almost within the range of common sense. Then it is clear that the electric force is everywhere doubled, and, further, that a unit test-charge would have twice as much work done on it during motion from the positive to the negative conductor ; that is to say, the potential difference between the conductors is doubled—it is $2V$ when the charges are each $2Q$. By similar reasoning it becomes evident that if the quantity of electricity on each conductor be increased m times the potential difference is increased m times. In other words, the ratio of the charge to the potential difference it causes is constant and always equal to Q/V. This ratio is called the capacitance of the condenser, or, frequently, the electric capacity. Giving it the symbol C we have

$$C = Q/V.$$

Energy of Charged Condenser.

17. This proportionality between charge and P.D. enables us to express in terms of Q and V the energy stored in the condenser, or, what is the same thing, the work that can be obtained by allowing the positive and negative charges to join and neutralise each other. This might be done, for instance, by connecting the conductors together through a very great resistance, so that a

small current flowed from high potential to low. The first portions of electricity to cross would pass at potential V, but when discharge has proceeded a quarter, say, portions of electricity crossing would have only $\frac{3}{4}V$ ergs of work done on each unit of electricity. Halfway through discharge a unit of electricity crossing would have $\frac{1}{2}V$ ergs of work done upon it, and this in fact would be the average amount of work done per unit during the whole process of discharge. Since, finally, Q units cross the work done is Q times the average potential difference, or

$$W = \tfrac{1}{2}QV.$$

This also

$$= \tfrac{1}{2}CV^2$$

$$= \tfrac{1}{2}Q^2/C.$$

If Q, V and C are in esu then W is in ergs; while if Q is in coulombs, V in volts and C in farads, W is in joules. The energy is spent, in the case supposed, in heating the material of the resistance.

As an example, let a condenser of 1 microfarad capacity be charged to 1,000 V. Then

$$W = \tfrac{1}{2} \,.\, 10^{-6} \,.\, 1{,}000 = \tfrac{1}{2} \,.\, 10^{-3} \text{ joule.}$$

A voltaic cell that has a terminal P.D. of 1 volt when 1 ampere is running would do as much work as this in one two-thousandth of a second.

18. Since the total energy of the condenser is $\frac{1}{2}QV$ and the number of Faraday lines is Q, the energy per Faraday line is $\frac{1}{2}V$. Of course each Faraday line is merely the representative of a bundle of lines of strain that start from a positive unit of electricity on one conductor and end on a negative unit on the other, and as the lines of strain cannot intersect, we see that each bundle keeps to its own region of space. We conclude that $\frac{1}{2}V$ units of energy associated with each Faraday line are stored in the region occupied by the bundle it represents, apparently as electric strain.

19. Hitherto we have supposed the region between the conductors to be vacuous. Let it now be filled with a dielectric of S.I.C. κ. Since the Faraday line density is unaltered the forces of the field will be reduced everywhere in the proportion κ; therefore the work done in passing a test-charge from one side to the other will be reduced in that proportion, that is, the P.D. will now be V/κ instead of V. It follows that the ratio of Q to

P.D. will now be κ times greater, or the capacitance will be κ times as great as in the case of vacuum. It was by this increase of capacitance that Faraday demonstrated the importance of the dielectric in the theory of the forces in the field.

Another consequence of the lowered value of V is that the energy stored by the condenser when charged with Q is less in that ratio κ; but if, on the other hand, the condenser be raised to its former potential difference by raising its charge to κQ, the energy stored will be κ times greater than before.

Special Forms of Condenser

We proceed to apply the general theorems obtained above to some particular cases, namely, the plate, the cylindrical, and the spherical condensers.

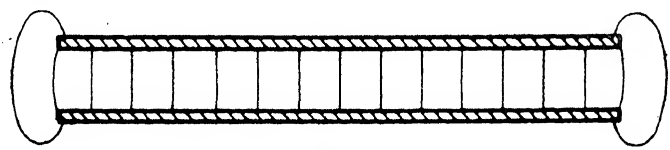


Fig. 13.

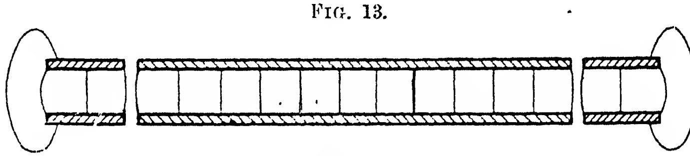


Fig. 14.

The Plate Condenser.

20. This is formed of two equal plates held parallel at a distance apart small compared with their linear dimensions. The field between them is uniform except near their edges, as indicated in Fig. 13. The uniform part can be selected, so to speak, by using "guard-rings," as in Fig. 14, which are touched momentarily to the plates after each charging or discharging. Let the area of each plate be A sq. cm (without guard-rings) and their distance apart be x. Let Q be the charge on each plate, V the potential difference, F the uniform field between the plates. The line density is here the same as the surface density of electricity, which is Q/A, and therefore

$$F = 4\pi Q/\kappa A.$$

But the work on taking unit charge across is

$$V = Fx = 4\pi Qx/\kappa A.$$

Hence

$$C = Q/V = \kappa A/4\pi x.$$

621.3842
N R1.1

3733

Cylindrical Condensers.

21. Take two coaxial circular cylinders of metal, as in Fig. 15, of radii x_1 and x_2, and of length a, all the lengths being in centimetres and a being much greater than x_2. It is obvious that the Faraday lines spread themselves radially in planes perpendicular to the axis except at the ends and that they form a cylindrical

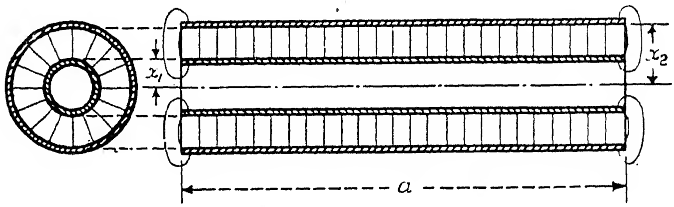


FIG. 15.

field in which each cylinder is an equipotential surface (see § 13). In consequence the difference of potential between the metal tubes is, neglecting the errors due to the ends,

$$V_{12} = (2q_1/\kappa) \log_\epsilon (x_2/x_1)$$

where q_1 is the number of Faraday lines per unit length of the condenser. The total charge on each tube is equal to the number of Faraday lines terminating on each, and this number is plainly the same for each, namely, aq_1. The capacitance C of the condenser is therefore

$$C = \tfrac{1}{2}\kappa a/\log_\epsilon (x_2/x_1).$$

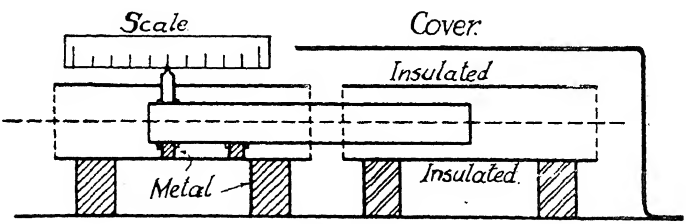


FIG. 16.

22. The tubular condenser forms one of the best variable standard condensers for the laboratory. For this its construction is modified, as indicated in Fig. 16. There are two equal fixed metal tubes, one carefully insulated; a tube sliding coaxially inside the fixed ones, and a metal cover. The insulated tube is one side of the condenser and the other tubes and cover, all

electrically connected, form the other side. The inner tube is balanced so that it does not require any mechanical support from the insulated tube. The change of capacitance due to a known

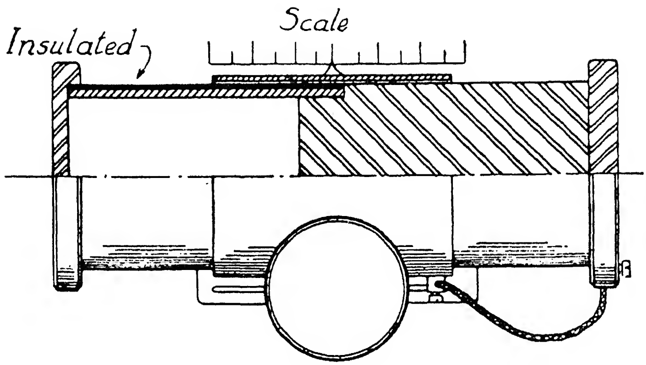


FIG. 17.

movement of the inner tube is given by the formula above with an accuracy limited only by the workmanship. It can be shown that a small proportional displacement of the axis of the inner tube from coincidence with the axis of the outer tubes has an even smaller proportional effect on the capacitance, which is a valuable feature in an instrument. Fig. 17 shows a form of cylindrical condenser used in wireless apparatus.

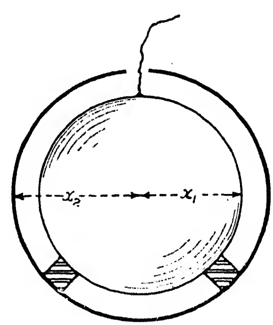


FIG. 18.

Spherical Condensers.

23. Two metal spheres of different sizes fixed concentrically constitute a spherical condenser. As indicated in Fig. 18, the field diverges uniformly in all directions, with, however, some disturbance at the place where the lead to the inner sphere passes through the outer one. Evidently the metal surfaces coincide with natural equipotential surfaces, such as those in Fig. 11, and we may at once state that

$$V_{12} = Q\left(\frac{1}{x_1} - \frac{1}{x_2}\right)$$

where Q is the total number of Faraday lines crossing the dielectric

and is therefore also the magnitude of the charge on each conductor. From this we obtain, for the capacitance,

$$C = Q/V_{12}$$
$$= 1 \Big/ \left(\frac{1}{x_1} - \frac{1}{x_2}\right)$$
$$= x_1 x_2/(x_2 - x_1).$$

It may be remarked that when x_2 is very large compared with x_1, we approximate to the state of things in which a sphere of radius x_1 cm is isolated in empty space, and we see that then its capacitance is

$$C_1 = x_1\,;$$

that is, its capacitance in electrostatic units is equal numerically to the number of centimetres in its radius. Hence the capacitance of any condenser or conductor in esu is often spoken of as so many "centimetres," meaning that the capacitance in question is equal to that of a sphere of radius so many centimetres in empty space. An antenna of capacitance 1,500 esu has the same capacitance as an isolated sphere 30 m in diameter.

Some Properties of Faraday Lines.

24. Returning to the plate condenser of Fig. 14 it is evident that the plates are pulled inwards by the tension of the Faraday lines. To estimate these forces we may calculate the work that would be done by the attraction if it were allowed to pull the plates together against some sort of mechanical resistance, that is to say, a brake. Let T_1 be the pull in each Faraday line, then the total pull is QT_1, and this is constant as x diminishes. Therefore the work done is

$$W = QT_1 x.$$

But this must equal the store of energy of the condenser, which is

$$W = \tfrac{1}{2}QV$$

Therefore

$$T_1 = \tfrac{1}{2}V/x$$
$$= \tfrac{1}{2}F = 2\pi\sigma/\kappa$$

where σ is the Faraday line density through the field. The fact that the tension in each Faraday line is proportional to the P.D. between its ends affords a logical justification for calling voltage "tension." A "high-tension battery" is one that puts each Faraday line between its terminals into a state of high tension. The pull per square centimetre in a bundle of lines of strain is

$$T = \tfrac{1}{2}F\sigma = 2\pi\sigma^2/\kappa$$
$$= \kappa F^2/8\pi.$$

The formulæ connecting F, σ, T_1 and T are true at any point of any field.

25. The whole energy stored in the electric strain between the plates is $\frac{1}{2}QV$, and therefore the amount stored in each Faraday line is $\frac{1}{2}V$, as already seen. The volume of the dielectric is Ax, and therefore the energy per unit volume is $\frac{1}{2}QV/Ax$, or $\frac{1}{2}\sigma F$ since $Q/A = \sigma$ and $V/x = F$. Thus the energy per unit volume has the same expression as the pull per square centi-

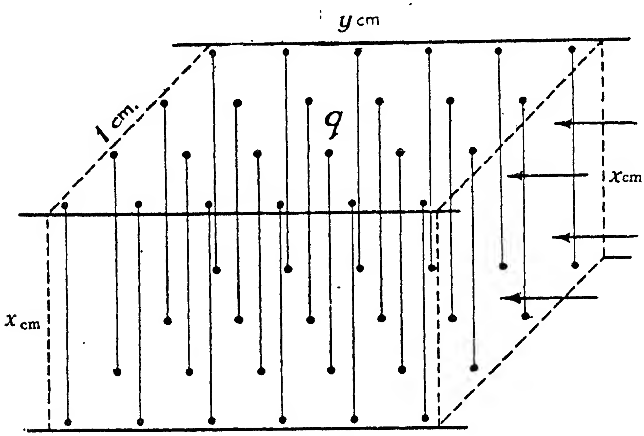


FIG. 19.

metre, as may easily be seen more directly. We have for the last expression

$$\text{energy per unit volume} = \kappa F^2/8\pi.$$

26. Perhaps the most striking property of the Faraday lines as brought to notice by the formulæ for the tension and the potential energy in them is that these are functions of the density of the lines at the place in question. In other words, the tension and energy of each line at any place, though not influenced by the length of the line, is increased by the presence of other lines. This dependence of the energy stored in each line on the closeness of the packing of the lines shows that work will have to be done in order to pack the lines in a condenser closer together than before; and shows in fact that if the areas of the plates be somehow diminished mechanical forces will arise to oppose the

attempted decrease. When a fixed condenser is charged by a voltaic cell the lines that crowd into the space between the plates have to be packed at the expense of the chemical energy of the cell. And thus it is logical to regard the charging of a fixed condenser as the act of forcing lines together in opposition to that mutual repulsion between them whose existence was perceived by Faraday and whose magnitude was evaluated by Maxwell. We are now in a position to repeat this evaluation, as follows:—

Let Fig. 19 represent a portion of the condenser taken from under the middle so as to avoid distortion of the lines, and let its dimensions be $x \times y \times 1$ in centimetres. Let q be the charge on this area, c the capacitance of this portion of the condenser, w the energy in the lines, and Y the mechanical force per unit area exerted by the neighbouring lines on those marked off by the diagram. We shall suppose the volume to yield a little by the shortening of y, so that only the mechanical force arrows doing work are indicated in the diagram.

The energy of the volume is

$$w = \tfrac{1}{2}q^2/c$$

and c is proportional to y. Thus the energy content is inversely proportional to y, or

$$w = m/y$$

where m is constant though y varies. Then when the volume is compressed laterally to the extent δy, we have an energy change given by

$$\delta w = \delta(m/y) = (m/y^2)\,(-\delta y).$$

The minus shows merely that increase of energy content is associated with decrease of y, when the charge on each plate is constant. But the force doing the work is $Y \times x \times 1$, and the work it does is also δw, therefore

$$\delta w = Yx(-\delta y).$$

Hence, on equating these two equivalents of δw, we obtain

$$Yx = m/y^2 = w/y$$

or

$$Y = w/xy.$$

But xy is the volume of dielectric, and therefore the right-hand side is the energy stored per unit volume. We know besides that the energy per unit volume is equal to the tension T (reckoned per square centimetre) along the Faraday lines. Hence

$$Y = T$$

Verbally, the lateral pressure is equal to the longitudinal tension; formulæ for this are given in § 21.

27. It is interesting to apply the above discussion to the well-known "vane" type of variable plate condenser. In this a number of parallel semicircular plates all connected to a spindle are carried in interlacing fashion between a set of parallel fixed plates by the rotation of the spindle by hand. Suppose such a condenser charged in the position of maximum capacitance and one side insulated. Then the turning of the spindle causes the lines to pack themselves closer; hence work is done mechanically. This is made evident electrically by the rise of the voltage between the plates.

28. The expression $\kappa F^2/8\pi$, which represents the energy density in the field, the tensional stress along, and the hydrostatic compressional stress across the Faraday lines, and, further, the expression $\frac{1}{2}F$, which represents the tension in a single line, all involve F the electric force in the medium, and are therefore all limited in magnitude by the fact that the electric force itself cannot be increased indefinitely. Dielectric media of every kind break down if the electric force in them is raised to a high value. The minimum force required to break a dielectric is different in different dielectrics and varies greatly with the circumstances in which the force is applied; it is spoken of as the electric strength, or the dielectric strength, of the material. For instance, in air under ordinary conditions the value of F cannot usually exceed about 100 C.G.S. electrostatic units, and therefore the mechanical tension in a Faraday line in air cannot greatly exceed 50 dynes. Again, the value of $\kappa F^2/8\pi$, under the same conditions, is about 400. Thus only about 400 ergs of energy can be stored electrically in 1 c.c. of ordinary air (or 0·84 ft.-lb. per cubic foot)—an amount that could be given to the same air in mechanical form by compressing it by 0·04 per cent. of its bulk. Since glass has about five times the strength of air and about five times the permittivity, its energy-storing capabilities are about 125 times as great as those of air at ordinary pressures and temperatures.

These formulæ connecting F, σ and the energy are true for a non-uniform field if F and σ are taken to be the values of the electric force and Faraday line density in an infinitesimal volume round the point of space concerned. Some applications of these results will appear later.

ELECTROMAGNETISM

29. The study of the electric current falls into two main divisions. The first is concerned principally with the transference of electricity, and embraces such subjects as cathode rays, the ionisation of gases, and conduction through electrolytes and metals; the second is concerned with the magnetic effects of currents and the propagation of waves. The second division is of immediate interest to us, and to this we now turn. We shall follow Oliver Heaviside's and J. J. Thomson's methods of combining the purely physical intuition of Faraday with the mathematical reasoning of Maxwell. For the purposes of our principal subject this is preferable to the orthodox way of discussing the magnetic effects of electric currents—which is based on Ampere's law that an electric current is in certain respects equivalent to a hypothetical distribution of magnetism—and, besides, it leads directly to an enlarged conception of electric current. For a fuller treatment of this point of view the reader is referred to J. J. Thomson's "Recent Researches" and later writings. Its essence consists in postulating that magnetic effects are due to the motion of Faraday lines, which, of course, spread throughout the space around the moving electricity to which they belong. As a preliminary it is necessary to review the salient facts of magnetostatics.

MAGNETOSTATICS

30. In magnetostatics the unit magnetic pole is defined as that which exerts on an equal pole at a distance of a centimetre in vacuum a mechanical force of 1 dyne, each pole being supposed concentrated at a point. This is precisely analogous to the definition of unit of electricity. Following the electric analogy, a magnetic pole is supposed to create at a distance x a magnetic force (or field, or intensity) given by

$$H = m/\mu x^2$$

where H is the magnetic force, m the magnitude of the pole, and μ a quantity characteristic of the medium called the permeability. The mechanical force exerted upon another pole of strength m' placed in the field of intensity H is $m'H$ dynes and is in the direction of the magnetic force. In consequence, by observing the direction in which a test-pole is urged in a given

magnetic field it is possible to map out the field by lines of magnetic force.

Magnetic Flux.

31. It has been found necessary, or at least convenient, to conceive that magnetic force produces wherever it acts a phenomenon called magnetic flux, or induction, formally analogous to electric strain; and the connection between force and flux density has been settled in accordance with the equation

$$B = \mu H$$

where B, the flux density, is the flux through the perpendicular square centimetre and μ is the permeability. The unit of flux density is fixed by this equation, and from the equation it is obvious that B corresponds to $4\pi\sigma$ in electrostatics (§ 7). If we represent the flux by lines analogous to the Faraday lines, the equation shows that we must suppose $4\pi m$ lines to be associated with a pole of size m. It is perhaps unfortunate that there is this numerical discrepancy between the reckoning of flux density and Faraday line density.

Since magnetic force obeys the same mathematical laws as electric force, the lines of magnetic force or flux have properties similar to those of electric force or strain. In order to avoid trouble with the discrepancy alluded to above in the reckoning of the flux and the strain it is well to express all the analysis in terms of the forces. We then immediately conclude that the magnetic lines may be regarded as having longitudinal tension, lateral pressure and volume energy expressible by

$$\mu H^2/8\pi.$$

This expression, if converted by introduction of B, becomes

$$BH/8\pi \text{ and } B^2/8\pi\mu.$$

Magnetic potential difference, or magnetomotive force (abbreviation mmf) is precisely analogous to electric potential difference or emf. In magnetostatics mmf between two points is the line integral of the magnetic force taken along any path joining the points, and inversely, the magnetic force at a point is the rate of decrease of the magnetic potential at that point and has the same direction.

Electromagnetic Units.

32. In the magnetic system of units, which is based on the definition of unit magnetic pole given above, the permeability of

a vacuum is arbitrarily assumed to be unity. It is not to be expected that this random assumption will fit with the electrostatic system of units, which involves the arbitrary assumption that the permittivity of a vacuum is unity; and indeed the two systems clash when the magnetic effects of electricity in motion come under investigation. In order to escape the dilemma and confusion arising from the conflict of the systems of units we shall develop our electromagnetic equations in a general manner, so that any system of units may be introduced into them at any stage; neither κ nor μ will be assumed unity in a vacuum (see § 99).

First Law of Electromagnetism

33. Electric current is usually conceived as consisting of a flow of electricity along a conductor; a current flows, for instance, along a wire connected between two plates of a charged condenser. Evidently in such a case there must be besides transference of electricity a movement of the lines of strain stretched between the plates of the condenser while it is charged, and since the magnetic effects discovered by Oersted appear in the space traversed by the moving Faraday lines it is natural to try to attribute the magnetic force to the motion of the lines of electric strain. This is found possible and leads to an extension of the meaning of "current."

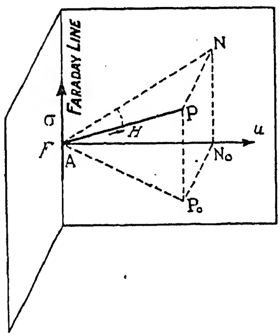


Fig. 20.

Moving Faraday lines are familiar in the following three phenomena:—

(1) Conduction currents.
(2) Convection currents.
(3) Dielectric currents and waves.

Of these the first is the most familiar and was the earliest investigated. The second kind of current appears when electricity is moved bodily through a dielectric, as happens when electrons pass along a vacuous tube in the form of cathode rays, or when positive or negative ions are blown through air, or when

a charged conductor is moved. Rowland and Hutchinson showed that if the two circular plates of an air condenser are rotated about an axle through their centres perpendicular to their planes, magnetic forces arise of the expected magnitude; and Röntgen showed that if the plates are stationary and a glass plate is introduced between them and rotated, magnetic effects are obtained. This raises the difficult question: Is the magnetic force due to relative motion between matter and the Faraday lines, and if so, what would occur in a perfect vacuum? No satisfactory answer appears to have been given yet to this question. Perhaps motion relative to the detecting apparatus is required, and perhaps Röntgen's result was due to surface charges of electricity on his glass.

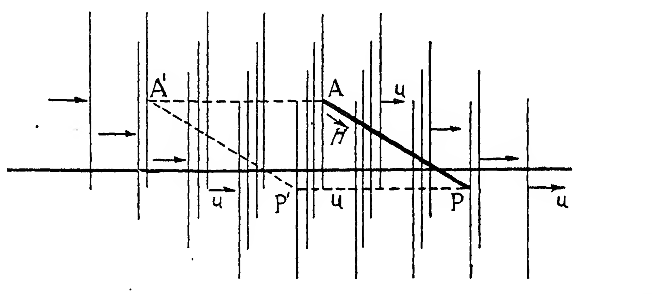


FIG. 21.

The third type of current, the dielectric current, is a conception due to Maxwell and a distinguishing feature of his theory of the electromagnetic field and of electric waves. It will be specially discussed below.

Statement of First Law.

34. The first law of electromagnetism may be stated thus: Faraday lines when cutting across an element of length AP produce along AP a magnetic force equal to 4π times the number of lines cutting AP per unit length per second. The permittivity and the permeability of the medium, if isotropic, have no effect. The sense of the magnetic force is shown in Fig. 20.

The first law is best made clear by a few illustrations of its application to an electric field of uniform intensity moving perpendicular to its Faraday lines with speed u cm per second.

First let AP be perpendicular to the lines and also to the direction of motion and let its length be a cm, as indicated in Fig. 21. Then the number of lines cutting AP per second is equal to the number passing through the rectangle APP′A′ at the instant defined by the figure, and this is $\sigma u a$ per second, or σu per second per unit length. Hence

$$H = 4\pi\sigma u$$
$$= \kappa F u.$$

Next let AP be inclined to the common perpendicular to σ and u, as in Fig. 20, where σ and u are in the plane of the paper and P is in front of that plane. Draw PN perpendicular to both σ and u, PP_0 parallel to σ, and P_0N_0 parallel to PN. Then the number of lines cutting AP per second is the same as the number cutting AP_0 or P_0N_0 or PN per second. This number is σu . PN; and therefore the number per second per unit length of the line AP multiplied by 4π is

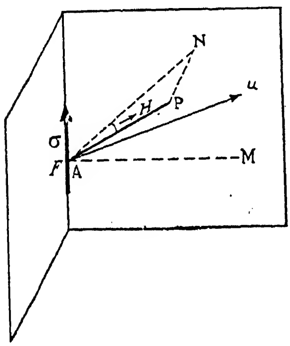


FIG. 22.

$$H = 4\pi\sigma u \,.\, \mathrm{PN/AP} = 4\pi\sigma u \sin \angle \mathrm{NAP}.$$

Note that the magnetic force is greatest in the direction perpendicular to both σ and u being

$$H' = 4\pi\sigma u.$$

Therefore $$H = H' \sin \angle \mathrm{NAP}.$$

35. Finally consider the general case where the lines are not moving perpendicularly to themselves, but in the direction indicated in Fig. 22. Here σ and u are, as before, in the plane of the paper, P is in front of that plane, and PN is perpendicular to it. Evidently the component of velocity perpendicular to the lines is the only motion carrying lines across AP, and this component may be written $u \sin (u, \sigma)$ where (u, σ) stands for "the angle between u and σ." Utilising the result of the last case we obtain now the result

$$H = 4\pi\sigma \sin \angle \mathrm{NAP} \,.\, u \sin (u, \sigma).$$

The above reasoning leads to another way of stating the first law: Moving Faraday lines produce a resultant magnetic field which is perpendicular both to themselves and to the direction

of their motion and is of intensity 4π times the product of their density and their velocity perpendicular to themselves; the magnetic field in any other direction is obtained by the usual methods of resolution in that direction. Though this has been deduced by aid of a uniform field it can be proved to hold at each point of a non-uniform field.

First Law in Terms of Mmf.

36. If the magnetic force generated near AP, and duly reckoned per unit length, be multiplied by the length of AP, we get the mmf being generated between A and P. The magnitude of this mmf is independent of the direction of the Faraday lines relative

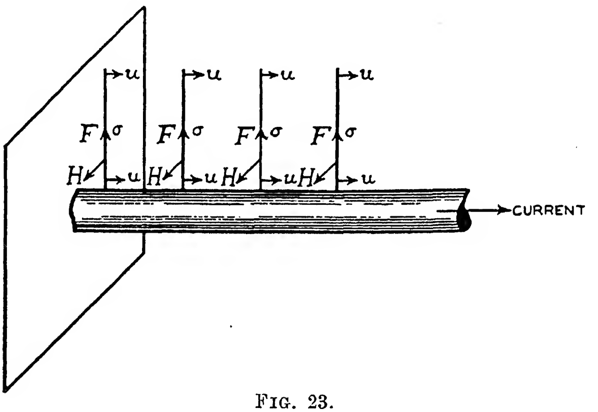


FIG. 23.

to AP, being determined solely by the number of lines crossing AP per second. We may join another element of length on to AP, say PC, and in fact build up any tortuous curve and say that along it, at any instant, the mmf is equal to 4π times the number of Faraday lines that cross the curve per second.

Rule for Direction of Magnetic Force.

37. Perhaps the easiest way of finding in any particular case the sense of the magnetic force due to the motion of a given set of Faraday lines is to imagine that the feet of the lines are sliding along a wire, as in Fig. 23. The direction of the magnetic field is then the same as that due to a current flowing in the wire in the direction of motion of the Faraday lines, that is to say, is deter-

mined by driving a right-handed screw along the wire and taking the sense of rotation as the sense of the magnetic field.

Circuital Form of First Law.

38. The first application we shall make of the last result is one that puts the first law into a well-known form. In Fig. 24 an electric field is represented by its Faraday lines, which are supposed to be moving in any general manner, and a closed curve of any shape has been drawn in the field. The curve is made up of elements of length in each of which magnetic force is generated according to the above law whenever Faraday lines cut them. Now lines cutting inward at any place will produce magnetic force in the same direction round the contour, namely, to the right hand in front, and lines cutting outward will produce force in the opposite direction; and we may conceive all these effects added together round the curve with due regard to sign. We shall then obtain the magnetomotive force round the curve, and we see that it must equal 4π times the excess of the number of lines entering the curve per second over those leaving it. We thus arrive at the circuital form of the first law:—

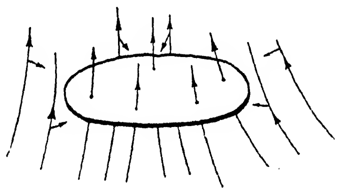


Fig. 24.

The magnetomotive force round a closed curve imagined drawn in any changing electric field (or the work done by the magnetic field on a unit pole guided round a closed curve) is equal to 4π times the rate of increase of the lines enclosed within the curve at the instant, provided that the curve does not embrace any line more than once. The sense of the mmf is related to the sense of the resultant electric strain in the same way as are the rotation and the direction of drive of a right-handed screw.

When the closed curve is reduced to an infinitesimal plane loop round a particular point in the field and turned so as to obtain the greatest possible mmf there at the instant, the mmf per unit area is called the "curl" or the "rotation" of the magnetic force at the point. The first law may then be stated: The curl of the magnetic force at a point is equal to 4π times the rate of increase of the electric strain there.

One advantage of stating the law in the circuital form is that

the difficulty presented by the idea of relative motion between an electric field and a vacuum, which arises in the first form of statement, is partially eluded.

Kinetic Energy of Faraday Lines.

39. Perhaps the most striking of the immediate consequences of the first law is that there exists in the field of moving Faraday lines a magnetic field H whose energy density must be $\mu H^2/8\pi$ as we have seen. That is to say, Faraday lines in motion have energy additional to their strain energy. Since in the uniform electric field of intensity F moving with perpendicular velocity u, the value of the resultant H is κFu, the magnetic energy is

$$\frac{\mu H^2}{8\pi} = \mu\kappa\left(\frac{\kappa F^2}{8\pi}\right)u^2,$$

which is proportional to u^2. The analogy with the kinetic energy of matter is unmistakable; the multiplier of u^2 is called the electromagnetic mass of the lines. It is noteworthy that if u^2 be made equal to $1/\mu\kappa$, the term on the right hand becomes the electric strain energy per unit volume; and the equation therefore shows that at the velocity $(\mu\kappa)^{-\frac{1}{2}}$ the kinetic and strain energies are equal. This velocity, as we shall see later, is the velocity of electric waves.

The magnetic effect in the field may, then, be regarded as being merely evidence of Faraday tubes being present and in motion. Many elaborate speculations have been based upon this idea, especially with regard to electromagnetic mass. Readers interested will find some of these discussed in J. J. Thomson's *Electricity and Matter.*

Transference of Energy.

40. Keeping still to the perpendicular motion of a uniform electric field let Fig. 25 represent a unit area in a fixed plane perpendicular to the direction of motion of the Faraday lines. At the instant pictured a rectangular prism of unit section and length u contains a quantity of Faraday lines and magnetic lines, and therefore a quantity of energy. At the end of one second all these lines and their energy will have swept through the unit area depicted. It is obvious that this amount of energy is

$$\left(\frac{\kappa F^2}{8\pi} + \frac{\mu H^2}{8\pi}\right)u.$$

We know that $$H = \kappa F u$$
and therefore the energy transmitted may be written
$$\frac{FH}{8\pi} + \frac{\mu H \cdot \kappa F u^2}{8\pi}.$$
or
$$\frac{FH}{8\pi}(1 + \kappa\mu u^2).$$
Let us suppose that $u = (\kappa\mu)^{-\frac{1}{2}}$; then the bracketed portion becomes 2. We thus arrive at a simple form of Poynting's Theorem: The energy carried through a unit area by Faraday lines parallel to the area and moving perpendicularly to the area with the velocity of electric waves is given by
$$FH/4\pi.$$

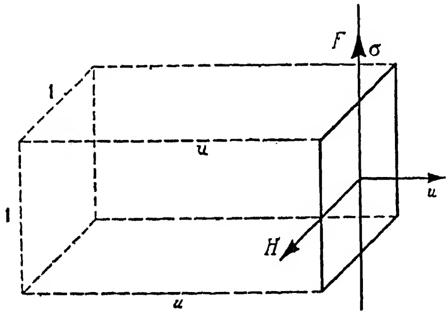


FIG. 25.

CONDUCTION CURRENTS

41. The magnitude of a conduction current along a wire is defined as the quantity of electricity passing a cross-section of the wire per second. The current density at a point in a conductor is the current reckoned per perpendicular square centimetre at

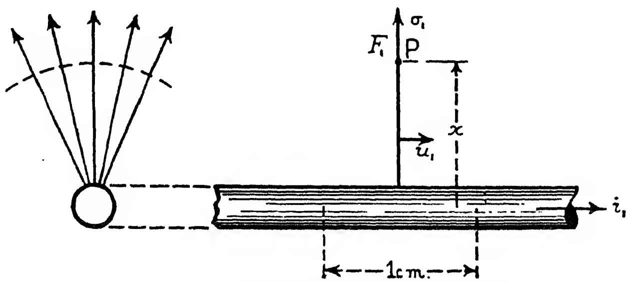


FIG. 26.

the point. A current may be pictured as a stream of electrons filtering through crowds of molecules, much larger than themselves, all chattering with the chaotic vibration we call heat; for

in conductors there appear to exist electrons that jump from molecule to molecule with enormous frequency and speed, enjoying during each journey a minute interval of freedom while passing from the rule of one molecule to that of another, and obeying during this interval the emf applied from outside to the conductor. If we suppose the net forward motion to be equivalent to that of q units of free electricity per centimetre length of wire at the speed u cm per second the current is $i = qu$. Usually it is indifferent to the magnetic effect whether we consider a large quantity to be moving at a slow speed or a small quantity at a high speed.

Field of Current in Straight Wire.

42. Let us consider the problem of a steady current along a straight round wire. We shall suppose the current to be accompanied by a procession of Faraday lines with their ends sliding along the wire. Since we do not know the inclination of the moving Faraday lines to the wire let us examine in turn two components of the lines, one set perpendicular to the wire, the other parallel to it, each set moving broadside-on to its own lines. These components are shown in Figs. 26 and 27. As regards the radial Faraday lines, if their velocity is u_1, they must number q_1 per centimetre length of the wire, equably distributed in all directions round the wire, and $i_1 = q_1 u_1$. At the point P distant x from the axis of the wire the line density $\sigma_1 = q_1/2\pi x$. Therefore the magnetic field at P is

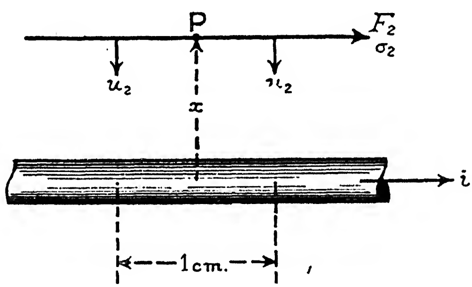


FIG. 27.

$$H_1 = 4\pi(q_1/2\pi x)u_1$$
$$= 2i_1/x.$$

This is a well-known formula for the magnetic force near the wire. The direction of the magnetic force at P is from the paper towards the reader. It is obviously of the same magnitude all round the circle of radius x.

43. As regards the parallel component Faraday lines in Fig. 27, to which is due the current i_2 along the wire, we must point out first that each Faraday line falling lengthwise on the wire collapses in the conducting material and causes the transference of one unit of electricity from left to right. It is plain that all the lines that so fall upon the wire must pass inwards through the circle of radius x, and if σ_2 is the density of the lines at distance x from the wire we must therefore have

$$2\pi x \,.\, \sigma_2 u_2 = i_2.$$

But the magnetic force at P is

$$H_2 = 4\pi\sigma_2 u_2$$

which

$$= 4\pi.\,(i_2/2\pi x)$$

from the last equation. Therefore

$$H_2 = 2i_2/x.$$

44. The radial and the parallel component lines thus both yield magnetic forces inversely proportional to the distance from the wire. Now when the component lines are compounded and their respective velocities also compounded, the resultant Faraday lines will make the same angle with the wire at all values of x, and the resultant velocity will everywhere be directed with equal inclination to the wire; this is indicated in Fig. 28. Hence the Faraday lines will be straight, will move forward and inward, and will produce magnetic force equal to the sum of H_1 and H_2, calculated above. If the total current be i and the total magnetic force H, we shall have

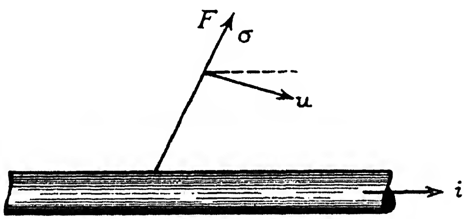


FIG. 28.

$$H = 2i/x$$

as may be confirmed by direct application of the first law to the sloping lines.

The motion of the sloping Faraday lines partly forward and partly inward has the effect of wearing away, so to speak, the feet of the Faraday lines. The portions thus destroyed in the conductor devote the energy they contain to warming the wire.

The energy of the joulean heat thus comes from outside the wire. The forward motion of the surviving portions of the Faraday lines carries the energy they contain to more distant parts of the conductor or to motors, say, connected to the wire.

45. Returning to the parallel component let us assume, on trial, that the Faraday lines move toward the wire with the velocity of electric waves and apply Poynting's Theorem (§ 40) to determine the energy flowing into the wire to generate joulean heat. At the surface of the wire we have, if its radius be r cm,

$$H = 2i/r.$$

The rate of flow of energy across each square centimetre of the surface of the wire is, by Poynting's Theorem, $FH/4\pi$. Therefore the total flow into a centimetre length of the wire is

$$\frac{FH}{4\pi} \cdot 2\pi r$$

or

$$F \cdot Hr/2$$

or

$$Fi.$$

Now, the electric force F at the surface of the wire and parallel to it is equal to the potential difference between the ends of the centimetre length, which may be written e. Thus the energy passed into the wire from the outside is

$$ei \text{ which } = Ri^2$$

if R is the resistance of the unit length of wire.

Equilibrium of Flux Lines in Field of Current.

46. It is conceivable that in a closed circuit formed of round parallel wire a state of things may arise similar to that of § 50 and only the magnetic field exist. The equilibrium of the magnetic lines is in this case of interest. Fig. 29 shows a half of a thin shell of space whose inner surface is at distance x from the centre of the wire, the magnetic field being $H = 2i/x$. The lateral magnetic pressures on the inside of the half shell exert a resultant force perpendicular to the base of the half shell and equal to $2xp$, where $p = \mu H^2/8\pi$. The difference of the inside and outside thrusts is $-d(2xp)$ outward. The question is: Is this differential outward thrust fully balanced by the hoop tension in the shell? The latter is $2pdx$, since the magnetic tension along the lines equals the magnetic lateral pressure and acts as shown. The question now is: Can

$$-d(2xp) = 2pdx$$

under the known law of p, given above, which may be written briefly

$$p \propto 1/x^2\ ?$$

On substituting we find that the equation is satisfied by this law, and therefore no other forces are needed to ensure equilibrium of the magnetic lines.

47. This result throws light on a view sometimes put forward that magnetic lines consist of or mark out an actual flow of a medium, a flow that possesses inertia in the sense that force of

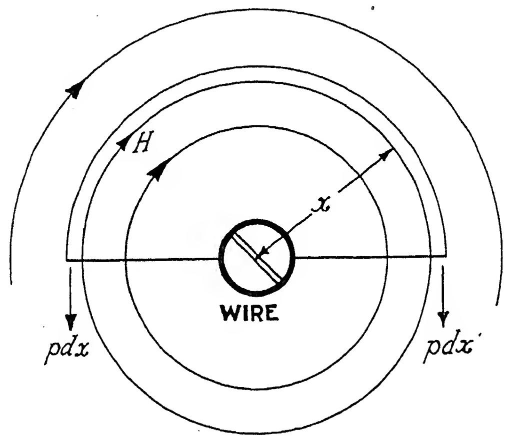


FIG. 29.

some kind has to be applied to alter its speed or the direction of its motion. Such a flow along circular lines in the magnetic field would give rise to a magnetic analogue of centrifugal force. The result just reached shows that no centrifugal force arises, and presumably that magnetic flux is not a mere flow of a medium along the lines of force. Instead, it seems more possible that the observed "inertia" of the magnetic lines is due to a spin of a medium round the lines as axes.

Fields of Currents along Conductors.

48. There are several simple ideal cases of flow of current that throw light on what probably occurs in real circuits. We shall consider first the current between two infinite parallel confronting plates lying horizontally and shall imagine Faraday lines stretched vertically between them to run from left to right and disappear.

Only very few lines are shown in the sketch (Fig. 30), but they must be supposed to be very numerous and of uniform density σ per square centimetre. If their velocity be u cm per second towards the right then the current along the lower plate is towards the right. That along the upper plate is one of negative electricity towards the right and must therefore be counted as equivalent to a leftward current of positive electricity. The magnitude of both is

$$\begin{aligned} i &= b\sigma u \\ &= \kappa F b u / 4\pi \end{aligned}$$

where F is the uniform electric force between the plates. The magnetic field produced by the motion of the lines is

$$H = \kappa F u = 4\pi i / b.$$

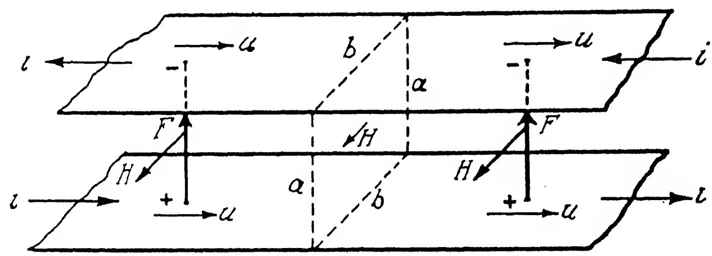


FIG. 30.

The electric tension, lateral pressure and energy density are

$$\kappa F^2/8\pi$$

and the magnetic tension, pressure and energy density are

$$\mu H^2/8\pi \text{ or } \mu\kappa^2 F^2 u^2/8\pi.$$

Now, the magnetic and the electric tensions, pressures and energy densities are equal if

$$\mu\kappa^2 F^2 u^2/8\pi = \kappa F^2/8\pi$$

that is, if

$$\mu\kappa u^2 = 1.$$

We shall see later that the value of u given by this equation is the velocity of electric waves, and that if we assume it to hold in this case it gives results consistent with experiment. Therefore we take for the magnetic and electric tensions, pressures and energy densities the same value, which may be called for short p. Thus

$$p = \kappa F^2/8\pi = \mu H^2/8\pi.$$

49. When the Faraday lines are moving at this particular speed the vertical lateral pressure between the magnetic lines

cancels the longitudinal tension of the electric lines, the tension of the magnetic lines cancels the lateral pressure of the electric lines, and the lateral pressures of both add in the direction of the motion of the Faraday lines. Now we have seen that the Faraday lines pull at the electricity in the plates and that the pull is transferred to the plates themselves; we shall see later that the magnetic pressures are similarly transferred to the plates by so-called electrodynamic force; hence we conclude that when the current is flowing the attraction that would be exerted between the plates if the electricity were stationary disappears. The sideways pressure likewise vanishes, but the pressure in the direction of motion is doubled.

Reflection of Faraday Lines.

50. Let the long parallel plates be now supposed terminated by a perfectly conducting perpendicular end plate, as indicated

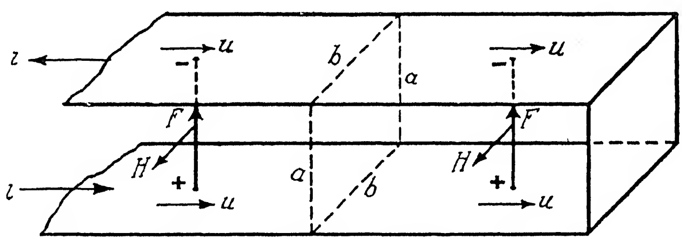


FIG. 31.

in Fig. 31. We now have a resistanceless go-and-return circuit. Any Faraday line reaching the end plate collapses into it; but it cannot disappear—its energy must be accounted for. Analogy with mechanical cases, especially the reflection of waves from an obstacle, suggest that the Faraday line is reflected turned upside down. This is found to fit in with observation, so we shall adopt the suggestion. Such an inverted line (not shown in the Figure) clearly produces magnetic force equal to and in the same direction as that it did before reflection. Moreover it adds to the current along the wires while cancelling a line of the original electric field. When a steady state has been reached all the electric field will be cancelled and only the magnetic field will be left between the plates. This appears to be actually the case in any good conducting circuit. Plainly the lateral repulsions of the magnetic lines are now unbalanced and their full force will be transferred

to the conductors. This electrodynamic force is referred to later.

Absorption of Faraday Lines.

51. Finally let the pair of long parallel plates be supposed terminated by an end plate of resisting material offering resistance R to the current. Faraday lines collapsing in this plate will spend their energy in warming it and will disappear. First let us suppose that this resistance is just the right size to absorb all the energy reaching it. Then the circumstances as regards the field between the plates will be exactly as in the case of the infinite plates already discussed. The potential difference between the plates must be Ri and also Fa, and since the value of i has been given above, we have

$$R = Fa/i = 4\pi(a/b)\sqrt{(\mu/\kappa)}$$

as the correct resistance for the end plate in order that there shall be complete absorption.

If the resistance of the plate is not just this value there will be reflections. If R is too small the circumstances will approach those of the non-resisting plate and there will be some reflection with reversal; while if R is too large there will be reflection without reversal. By assigning a symbol to the reflected portion for any given value of R and expressing that the new field must give the same potential drop as the current does in the resistance, the reader will be able to calculate the reflection.

Magnetic Field in Solenoids

Straight Solenoid.

52. The circuital form of the first law may be usefully applied to the most general kinds of conduction current. As an example,

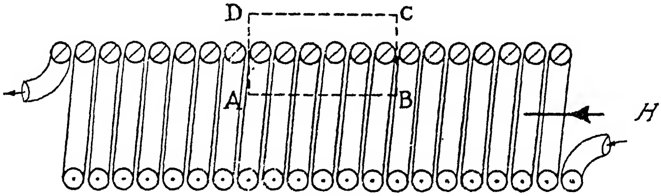


Fig. 32.

consider a closely-wound cylindrical coil or solenoid of length much greater than the diameter and let us apply the law to the

imaginary curve ABCD shown in Fig. 32 while a current i traverses the winding. Let the number of turns per unit length be τ. Symmetry shows that the magnetic field will be parallel to the axis of the cylinder, and it is clear that the field outside the coil will be very weak if the coil is very long. In order to measure the mmf we guide a unit magnetic pole round the curve ABCD and reckon up the work done by the magnetic field. This magnetic work is H . AB along AB, zero along BC and DA because there is no component field perpendicular to the coil, and negligible along CD as has been already stated.

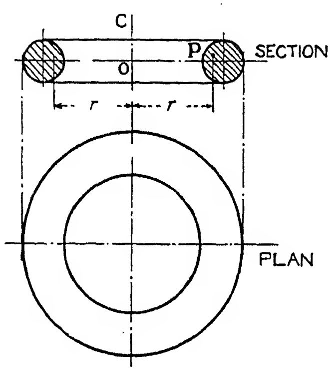


FIG. 33.

This by the first law is to be equated to 4π times the current surrounded by the curve, which is AB . τi. Thus for a straight solenoid

$$H \,.\, \mathrm{AB} = 4\pi \,.\, \mathrm{AB} \,.\, \tau i$$

or

$$H = 4\pi\tau i.$$

The field is of this intensity at all distances from the axis of the cylinder within the windings. Near the ends of a solenoid the field falls off in strength because of the splaying of the magnetic lines and the calculation of the field there needs advanced methods.

Ring Solenoid.

53. Let the ring solenoid, as it is called, of Fig. 33 have τ close and equally spaced turns in its circumference, each of them in a plane passing approximately through the axis of the tore. Take any circle centred on the axis, of radius CP $= r$, as shown in the sectional elevation, and imagine a unit pole guided round the circumference. The field at all points of this circle will be the same, by symmetry; let it be H. Thus the first law gives

$$H \,.\, 2\pi r = 4\pi\tau i$$

or

$$H = 2\tau i/r.$$

Dielectric Currents

54. When a condenser is being charged by aid, say, of a battery of voltaic cells there is, in the language of the two-fluid theory, a current of positive electricity flowing along the wire connecting the positive pole of the battery to one plate of the condenser, and an equal current of negative electricity flowing towards the other plate. The latter current is, for all our purposes, equivalent to a positive current flowing away from the plate concerned. Hence, in Fig. 34 the currents in the connecting wires are marked in the same sense round the circuit. Of course the current is but a transient one and as charging proceeds the backward potential difference of the condenser rises till it balances the emf of the battery, and then the current stops.

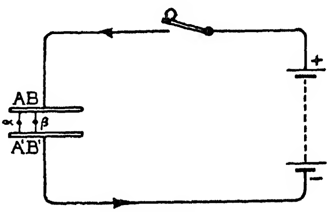


FIG. 34.

During the period of time when the current is flowing the charges on corresponding unit areas AB, A′B′ far from the edges of the plates continue to grow and more and more Faraday lines crowd into any area such as $\alpha\beta$ drawn in the dielectric between and parallel to the areas AB, A′B′.

In the first place it is plain that positive electricity can enter the area AB and negative electricity the area A′B′ only by flowing along the plates and crossing the boundaries of the areas. In the second place the Faraday lines, which accompany the charges, can crowd into the area $\alpha\beta$ only by moving sideways across its boundary. Let N be the number of lines through $\alpha\beta$ at any time, then N is also the quantity of positive electricity on AB and of negative on A′B′ at the same instant. It follows that the rate of increase of charge on AB, which is in other words the current flowing into AB, equals the rate of increase of the Faraday lines through $\alpha\beta$, which is in symbols dN/dt. The quantity dN/dt is called the *dielectric current* through $\alpha\beta$. Usually the area $\alpha\beta$ is understood to be of unit area and then the dielectric current at a place is the rate of increase of the density of the Faraday lines there, or $d\sigma/dt$. In all the cases with which we shall have to deal the direction of the Faraday lines (and of the electric force)

remains the same at any point even while their density is changing; in such cases the direction of the dielectric current at any point is defined as being the same as that of the Faraday lines if the density of these is increasing, but opposite to that of the Faraday lines if the density is diminishing.

55. Taking any form of condenser connected into a circuit, it is evident that at every instant the positive electricity on one side is linked with the negative on the other by Faraday lines passing through the dielectric. The whole dielectric current between the plates is equal to the rate of change of these lines at the instant and thus to the rate of change of either charge. We may therefore say that a transient current in the circuit of Fig. 34 is partly a conduction current and partly an equal dielectric current, or we may say that the current round the circuit is closed by the dielectric current. This conception is due to Maxwell, who spoke of such currents as displacement currents. As already mentioned, the dielectric strain may indeed consist in an actual mechanical displacement and elastic separation of the positive and negative portions of the molecules of the insulator. The term seems to imply that such displacements occur also in a vacuum, and therefore it is perhaps better to use the other term.

When we apply the first law to an element of length in the field of a pair of condenser plates we obtain the same result as would be obtained on applying Ampere's formula (§ 57) to the conduction currents running in the plates and leads. But dielectric currents occur in free electric waves where the application of Ampere's formula would be inappropriate, and then the conception of dielectric currents is invaluable.

Convection Currents

Magnetic Field of a Moving Point Charge.

56. We shall suppose the velocity u of the charge q to be much smaller than that of electric waves and shall find the magnetic force at the point P distant r from q. Symmetry shows that Faraday lines on the cone of angle θ with vertex at q is cutting the circle of radius y everywhere at this instant with the same speed and producing equal magnetic

force at each point of the circle. Let this be H. The Faraday line density at P is $q/4\pi r^2$, its component perpendicular to the direction of motion is $q \sin\theta/4\pi r^2$, and therefore by the first law

$$H = qu \sin\theta/r^2 = quy/r^3.$$

A right-handed screw driven along the direction of u gives the sense of H (see Fig. 35).

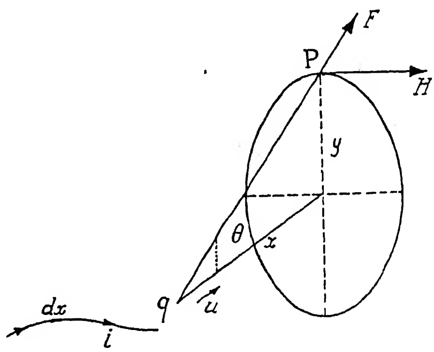


FIG. 35.

AMPERE'S FORMULA

57. We have seen in § 44, II., that whatever may actually be the direction of the Faraday lines associated with an electric current along a wire, the magnetic field deduced by aid of the first law on the supposition that the current is a slow flow of electricity along the surface of the wire agrees with the observed facts of the fields of currents. Hence we may venture, since rigour is not our aim here, to assume that the result just obtained will hold good for an element dx of a thin conductor carrying a current i if we replace qu by $i \,.\, dx$. We thus obtain the celebrated formula of Laplace and Ampere for the magnetic field due to a current element (dx in Fig. 35), namely

$$dH = idx \sin\theta/r^2 = idx \,.\, y/r^3,$$

where dH is written instead of H because it is only a contribution to the whole magnetic field of the complete current circuit.

Magnetic Force at a Point in the Axis of a Circular Current.

58. Since in Fig. 36 the point P on the axis is equidistant from every point on the circle the contribution from each element of current will be equal to

$$ids/r^2.$$

The direction of the force will, however, be different for each

element; it is indicated at PM for a typical element. But the component of PM perpendicular to the axis will be annulled by that due to a current element at the other end of the diameter of the circle, and therefore we attend only to the component along the axis. This will be for any one element

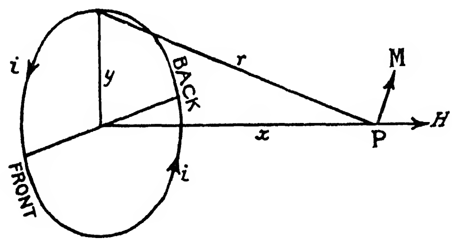


FIG. 36.

$$ids/r^2 \times y/r.$$

The addition of the effects of all the elements is simple, since i, y and r are constant; for the sum of ds is the circumference, that is, $2\pi y$. Thus

$$H = 2\pi i y^2/r^3.$$

For the particular case when P is at the centre of the circle, we have $r = y$ and then

$$H = 2\pi i/y.$$

Energy in the Field of a Current

Self and Mutual Inductance

59. Let the circuit depicted in Fig. 37 be of fine wire and be carrying a steady current i; loops of magnetic flux will thread it in number proportional to the current. One of these loops is selected for attention and the bundle of flux of which it is the representative is sketched round it. In any such space the energy stored per cubic centimetre is $\mu H^2/8\pi$ or $BH/8\pi$. Now if α is the sectional area in square centimetres of this bundle belonging to one loop of flux we must have $\alpha B = 1$, since whatever our system of units one loop is the unit of flux. Thus the energy $\alpha BH/8\pi$ of 1 cm of a flux loop becomes $H/8\pi$ and of x cm is $xH/8\pi$. In words, the energy in any portion of a magnetic loop is equal to the mmf between the

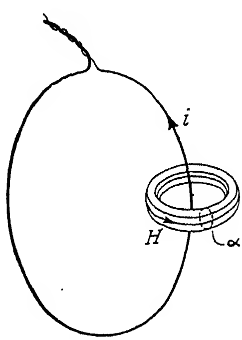


FIG. 37.

ends of that portion divided by 8π. Now the mmf in a whole loop of flux, whatever its shape or size, is $4\pi i$ according to the first law of electromagnetism; and therefore the energy associated with one complete flux loop encircling the current to which it is due is $4\pi i \div 8\pi$ or $\frac{1}{2}i$. If we represent the number of loops of flux through the circuit by the symbol $\mathscr{F}$ then the whole energy of the magnetic field of the current is

$$\tfrac{1}{2}\,\mathscr{F} i.$$

This is sometimes called the electrokinetic energy of the current.

60. In Fig. 38 a single circuit made of two turns of wire is depicted and with it are shown linked several magnetic loops—two small, two of medium size and two large. These loops are sketched as if the loops belonging to the separate turns maintained their identity, though as a fact they run together and form complicated closed curves linking with one or both turns of wire. The closest experimental scrutiny shows that all their electrical effects are the same whether they are compounded or not. Our method of sketching them shows that the turn marked a has more and more flux through it from turn b the closer b comes to a, and that this is over and above its own flux. Let $\mathscr{F}_a$ represent the amount of flux through a when b is far away, $\mathscr{F}_b$ the amount through b under the same conditions, $\mathscr{F}_{ba}$ the amount of b flux through turn a when the turns are brought near together as in the Figure, and $\mathscr{F}_{ab}$ the amount of a flux through b in the Figure. Therefore in the position shown in the Figure, the energy in the field of the current is

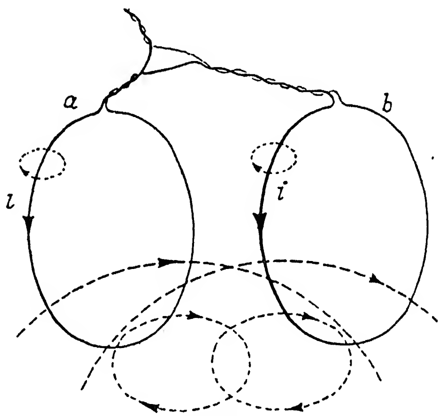


FIG. 38.

$$\tfrac{1}{2}\,\mathscr{F}_a i + \tfrac{1}{2}\,\mathscr{F}_{ba} i \qquad \text{due to flux through } a$$

and

$$\tfrac{1}{2}\,\mathscr{F}_a i + \tfrac{1}{2}\,\mathscr{F}_{ab} i \qquad \text{due to flux through } b.$$

The total is

$$\tfrac{1}{2}(\mathscr{F}_a + \mathscr{F}_b + \mathscr{F}_{ba} + \mathscr{F}_{ab})i.$$

The part inside the brackets is called the "flux linkages" of the whole circuit comprising turns a and b. It is not the same thing as the flux through the whole circuit, for in reckoning it we have counted certain loops twice because they linked with both turns.

Self-Inductance

61. The flux linkages increase in number when the current increases because the flux density at every point of space is proportional to the current. This linear proportionality between linkage and current enables us to write the flux linkages as

$$Li$$

where L stands for the linkages produced by the unit current. Evidently it is a factor that depends on the size, shape and relative positions of the parts of a circuit and on the law governing the spread of flux through space. This electro-geometrical quantity is called the "self-inductance," or the "inductance," or the "coefficient of self-induction" of the circuit. By § 60 we may write the energy of the field of the current i in the form

$$\tfrac{1}{2}Li^2.$$

62. If we like to consider separately the turns making up the circuit of Fig. 38 we may write, instead of the last expression,

$$\tfrac{1}{2}(L_a + L_b + L_{ba} + L_{ab})i^2$$

which conveys the information that the whole inductance of a circuit is not necessarily equal to the sum of the inductances of its parts. In Fig. 38 the turns are represented in a position where the flux from b through a threads the a turn in the positive direction, that is, the right-handed screw direction with regard to the a current; and similarly for the other mutual flux. Therefore $\mathscr{F}_{ba}$ and L_{ba}, $\mathscr{F}_{ab}$ and L_{ab}, are all positive. But if one turn were rotated about a diameter through 180 degrees all these would become negative. Thus L may be greater than, equal to or less than $L_a + L_b$. Extreme cases appear if the b coil is made exactly equal to and moved into coincidence with the a coil; for then, obviously, the linkages L_a, L_b, L_{ab}, L_{ba} all become equal and

$$L = 4L_a \text{ or zero}$$

according as the mutual linkage is positive or negative.

Mutual Inductance

63. Whatever its value the mutual linkage produced by unit current is called the "mutual inductance." It will be apparent later that unless there are dielectric currents distributed along the windings, the linkages that a makes with b are equal in number and sign to those that b makes with a when the same current flows through both turns. Therefore it is usual to give the mutual inductance L_{ab} and L_{ba} the same symbol M. For the simple case of Fig. 38 we may write for the total self-inductance

$$L = L_a + L_b + 2M$$

where M may be positive or negative.

One important conclusion which can be drawn immediately from the above reasoning is that in complicated circuits consisting of several turns or nearly closed portions the total self-inductance is equal to the sum of the self-inductances of the parts together with the mutual inductance of each part with every other part. This is a principle frequently applied in the calculation of the self-inductance of coils.

Energy of Two Current Fields.

64. In § 62 the turns of wire have been supposed connected in series so that the same current runs through both; but they may be separately served with equal currents without affecting the reasoning. In Fig. 39 the results are extended to the case of two turns of wire carrying different currents i_a and i_b. The flux through turn a is $L_a i_a$ due to its current i_a, and Mi_b due to the current i_b; the flux through turn b is $L_b i_b$ and Mi_a. Therefore the energy in the loops linking with turn a is

$$\tfrac{1}{2}L_a i_a{}^2 + \tfrac{1}{2}Mi_b i_a$$

and that in the flux loops through turn b is

$$\tfrac{1}{2}L_b i_b{}^2 + \tfrac{1}{2}Mi_a i_b.$$

The total energy in the field of the two currents is

$$\tfrac{1}{2}(L_a i_a{}^2 + L_b i_b{}^2 + 2Mi_a i_b),$$

where M may be positive or negative, but is shown positive in the figure.

Suppose now that the single turn a in Fig. 39 be replaced by a circuit of two turns like that of Fig. 38 and the single turn b similarly replaced. We have already seen that we may express the energy in the field of the first complete circuit in the form

$\frac{1}{2}L_1 i_1^2$ if i_1 be the current traversing it and L_1 the linkages created by unit current. Similarly the energy of the second complete circuit may be written $\frac{1}{2}L_2 i_2^2$ if i_2 be the current through this circuit. Of the flux produced by i_1 the amount Mi_1 will link with the second circuit and of that due to i_2 the amount Mi_2 will link with the first circuit. Therefore for this pair of two-turn circuits the total energy is

$$\tfrac{1}{2}(L_1 i_1^2 + L_2 i_2^2 + 2Mi_1 i_2).$$

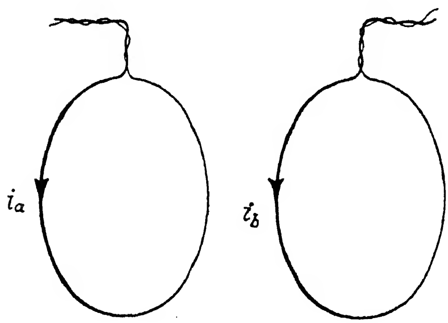


FIG. 39.

It is evident that the same rules will hold for all circuits, however complicated, provided that L_1, the self-inductance of the one, is defined as the flux linkages with circuit 1 when unit current traverses it, and L_2 a similar quantity for the second circuit; while M is the flux linkages made with one circuit due to unit current in the other. Any expression, such as that above, for a quantity of energy must yield positive values whatever magnitudes, positive or negative, the variables may take. The condition that the above expression shall be positive for all values of i_1 and i_2, is, by a well-known algebraic theorem,

$$M^2 \ngtr L_1 L_2.$$

Theorem of Mutual Linkages.

65. The important theorem just alluded to, that the linkages made with any circuit by unit current in another are equal in number to those made with the latter circuit by unit current in the former, may be proved in the following manner. In Fig. 40 AB represents a current element of one circuit and PQ any current element of the other circuit. The line joining AP is of length r and has been turned so as to lie in the plane of the paper; AB is also in the plane of the paper and therefore in general PQ stands out from the paper. The elements are drawn vastly out of proportion so that they shall be visible. The element PQ is shown resolved into three mutually perpendicular components of

which PN is parallel to AB. One of the rings of flux due to unit current in AB is also indicated.

The process we are going to use consists in imagining the current element AB to be brought up from a great distance by a movement perpendicular to itself from left to right. Evidently during this movement its magnetic rings will cut through PQ and enter the circuit of which PQ forms an infinitesimal portion. But inspection shows that the number cutting through PQ is equal to the number cutting through PN, since the motion cannot cause any rings to cut across QR or RN. Hence the motion up

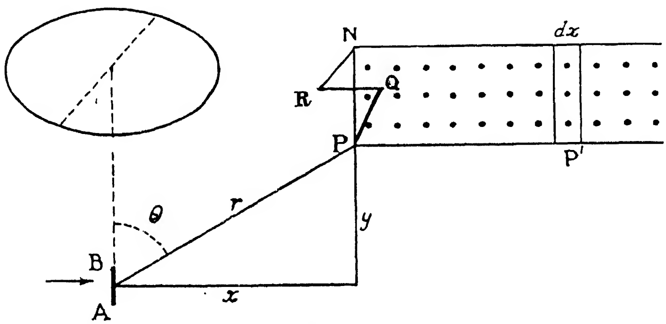


FIG. 40.

to the present position results in passing flux through PQ to the amount now threading the strip of space shown in the figure reaching from PN to infinity. We have to count the flux rings linked with this strip.

First we see by Ampere's law (§ 57) that the flux density produced at P by unit current in AB is

$$\frac{\mu \cdot \mathrm{AB} \cdot \sin\theta}{r^2} \quad \text{or} \quad \frac{\mu \cdot \mathrm{AB} \cdot x}{r^3}.$$

If we allow x and r to relate not only to the point P but also to any point such as P′ the same formula holds good, of course. Consequently the flux threading an infinitesimal length dx of the strip is PN . dx × flux density. Add all such quantities up by integration from the value of x shown in the figure to $x = \infty$ and get

$$\mathrm{AB} \cdot \mathrm{PN} \cdot \mu \int_x^\infty \frac{x}{r^3}\, dx$$

We shall represent the angle NPQ by ϵ, and put $x^2 + y^2$ for r^2 and so obtain

$$\mathrm{AB} \,.\, \mathrm{PQ} \cos\epsilon \,.\, \mu \int_x^{\infty} \frac{x\,dx}{(x^2 + y^2)^{3/2}}.$$

The integral becomes, noticing that y is constant,

$$\left[\frac{-1}{(x^2 + y^2)^{\frac{1}{2}}}\right]_x^{\infty} = \frac{1}{r}.$$

Hence the flux that has cut through PQ during the motion of AB up to the position shown is

$$\mathrm{AB} \,.\, \mathrm{PQ} \,.\, \cos\epsilon \; \mu/r.$$

Now PQ is any element whatever of a closed circuit (not drawn in the figure), so we may write $\mathrm{PQ} = ds$. Clearly every magnetic ring linking with the circuit must have entered the circuit by cutting through the wire at some point or other during the advance of AB to its present position, and those linking with several turns must have cut through the wire as many times as there are linkages. Some of these entries may have been made in the positive direction, some in the negative; but that can all be expressed by the language of the integral calculus. In short, we must integrate with respect to s all round the circuit, round every turn and bend, in order to obtain the total flux linkages made with the circuit by unit current in AB. This may be very difficult to carry out but is easily expressed symbolically in the form

$$\mu \mathrm{AB} \int \frac{\cos\epsilon}{r}\, ds.$$

The angle ϵ varies from point to point, in general, as well as r; and as the integration proceeds round the circumference the cosine will sometimes be positive, sometimes negative. It is important to emphasise that this integral gives the linkages with, and not merely the flux through, the circuit, and that it holds good for any position of AB whatever the size of AB or its inclination to r.

66. Now imagine that other elements carrying current alike AB are brought up and put end to end so as to form with AB a com-

plete circuit in which unit current is flowing; and suppose the linkages of each of these with the other circuit computed as described above, and finally all the results added together. This is expressed by writing ds' for AB and introducing an additional integral sign. Thus we obtain for the mutual linkages of the two circuits the equation

$$M = \mu \int \int \frac{\cos \epsilon \,.\, ds ds'}{r},$$

where ϵ is the angle, r the distance, between any two elements ds and ds' of the circuits.

The symmetry of this result shows that if we suppose the current to run in element ds instead of in element ds' we shall obtain the same value for M. This establishes the statement made in anticipation in § 63.

Throughout the above reasoning the circuits have been supposed made of wire of infinitesimal diameter, that is, the thickness of the wires has been neglected. It can be shown that thickness introduces no error when the current is symmetrically distributed round the centre line of the wire and this line is taken to represent the circuit in the calculations.

Application to Self-Inductance.

67. The self-inductance of a wire of finite section can be regarded as the mutual inductance of two equal filamentary circuits parallel to the centre line of the wire and at a distance apart equal to what is called the geometric mean distance of the section. The proof of this theorem and the calculation of the geometric mean distances for various sections of wire will be found in Maxwell's treatise. For a round wire the geometric mean distance is 0·7788 of the radius.

Calculation of Inductance.

The accurate calculation of formulæ for the inductance of circuits of even simple geometrical forms is often difficult and is in nearly all cases beyond the range of the mathematics assumed for the purposes of this book. There are a few instances, however, which will serve as illustrations of the principles discussed in the preceding paragraphs.

Induction of Solenoids.

68. Referring to Fig. 32, p. 51, in a unit length of the winding there are τ turns and the flux through each turn is

$$\mu HA \text{ or } \mu \,.\, 4\pi\tau A$$

when unit current flows, if A is the sectional area of the core of the winding. The linkages per unit current per centimetre run of the solenoid amount therefore to

$$L = 4\pi\tau^2\mu A.$$

If the core is of circular section radius a, so that $A = \pi a^2$, we see that

$$\begin{aligned} L &= 4\pi^2 a^2 \tau^2 \mu \\ &= (2\pi a\tau)^2\mu \\ &= l^2\mu \end{aligned}$$

where l is the length of wire wound upon the 1 cm of solenoid.

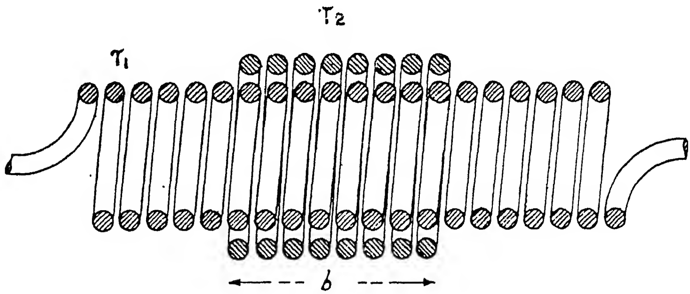


FIG. 41.

It will be noticed that in this computation the linkages made between current and flux inside the copper are neglected, that is to say, the wire is supposed infinitely thin. The correction for this appears in § 70.

The total inductance of a ring solenoid is easily calculated in the same manner, that is by calculating the flux through a section of the core and multiplying by the total number of turns; as an illustration it adds nothing to the above.

69. The mutual inductance between portions of two solenoids is easily calculated in the case described by Fig. 41. One coil is a fairly long solenoid with τ_1 turns per centimetre run, the other is a short coil of τ_2 turns per centimetre run, placed well away from the ends of the first. All the flux in the core of the

first, and no more, passes through the outer coil also. Using the same symbols as before, we have for the mutual linkages

$$M = BA \cdot \tau_2 b$$
$$= 4\pi\mu\tau_1\tau_2 bA.$$

The piece of apparatus represented by Fig. 41 forms a useful laboratory standard of mutual inductance. Note that if the value of μ be taken on the assumption that the permeability of a vacuum is unity, the result is expressed in absolute electromagnetic units.

Linkages inside a Round Wire.

70. When current flows along a wire a magnetic field arises inside the wire as well as outside and the first law of electromagnetism and the deductions we have made from it in § 59 apply. We may compute the contribution of this internal field to the self-inductance of a wire as follows:

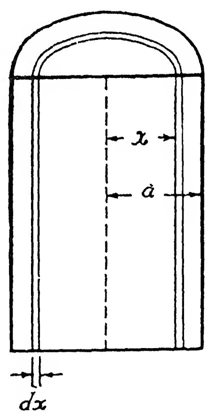


FIG. 42.

Fig. 42 shows one-half of a unit length of the round wire and a shell of copper of radius x and thickness dx is displayed. The flux loops in this shell are, by symmetry, circles of radius x approximately, and their number has to be determined by calculating the field H in the shell. Let the current along the wire be J; then the current flowing inside the shell is $(x^2/a^2)J$, since sectional areas are proportional to the squares of radii. The first law now gives

$$2\pi x H = 4\pi(x^2/a^2)J$$

or

$$H = 2Jx/a^2.$$

The flux in the shell is Bdx and this links with the current inside the shell, namely, Jx^2/a^2. Therefore the flux linkages for this shell are

$$Bdx \cdot Jx^2/a^2 \text{ or } (2J^2\mu x^3/a^4)dx.$$

This is true for any shell between the centre line and the surface

of the wire. Add up for all shells from the centre to the surface and obtain

$$\frac{2J^2\mu}{a^4}\int_0^a x^3dx = \frac{2J^2\mu}{a^4}\cdot\frac{a^4}{4}.$$

Hence $$L = \tfrac{1}{2}\mu$$

is the inductance of the wire per unit length so far as its internal field is concerned. It is evident from this result that an iron wire may have hundreds of times the internal inductance of a copper wire. The internal inductance depends in fact on the material and not on the diameter of the wire. On the other hand, it will be found on reference to books giving the formulæ, or to § 71, that the inductance due to the field outside the wire, which is usually the chief part of the inductance, depends greatly on the diameter and not at all upon the material of the wire. The smaller the diameter the greater is the external inductance.

Approximate Calculation of Inductance.

Two or three examples of methods of obtaining approximate formulæ will be given here because they suggest modes of solution of special problems that may present themselves on occasion to the reader.

Inductance of a Circle of Round Wire.

71. We have to estimate the flux threading the circle, and it is therefore first necessary to know the flux density at any point

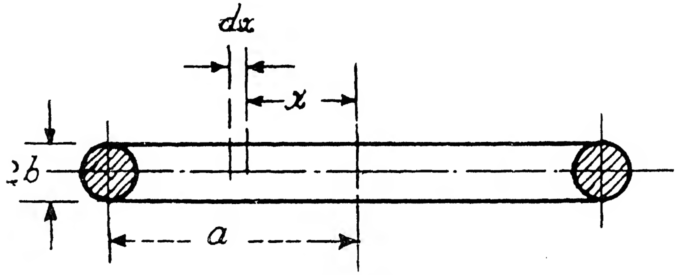


FIG. 43.

distant x from the centre of the circle. Then we take a narrow annulus of width dx and of radius x, as indicated in Fig. 43, and

write the flux through it as equal to circumference of annulus $\times$ width of annulus $\times$ flux density,

or
$$2\pi x dx \,.\, B$$
or
$$2\pi\mu H x dx.$$

We know that at the centre of the circle (§ 58) the field is $2\pi/a$ when the current is unity, where a is the radius of the circle; and that the field very near the wire is approximately equal to that which would be produced by a straight wire if we suppose the radius of the wire very small compared with the radius of the circle—that is, the field near the surface of the wire outside it is $2/(a - x)$ approximately. We now have to guess at a simple formula for the value of H at distances from the centre between these two extremes, and might start out by assuming

$$H = \frac{2\pi}{a} + \text{a multiple of } \frac{2}{a - x}.$$

The multiplier required must contain x in such a way that when $x = 0$ the value of H reduces to $2\pi/a$ and when $x = a - b$ it reduces to $2/b$. The very simplest multiplier that satisfies the former condition is mx, where m is independent of x. Try it in the equation for H, put $x = a - b$ and solve for m; we get

$$m = \frac{1}{a} - \frac{\pi b}{(a - b)a}.$$

As a rule b/a is less than 0·01, and therefore we may take for our approximate purpose $m = 1/a$, and then

$$H = \frac{2\pi}{a} + \frac{x}{a} \cdot \frac{2}{a - x}.$$

This rough value gives for the flux through the elementary ring of radius x the expression

$$\frac{4\pi\mu}{a}\left(\pi + \frac{x}{a - x}\right) x dx$$
or
$$\frac{4\pi\mu}{a}\left\{\frac{a^2}{a - x} + (\pi - 1)x - a\right\} dx.$$

Integrate this from $x = 0$ to $x = a - b$ and we obtain for the total flux through the circle (due to unit current in it)

$$4\pi\mu a\left\{\log_e \frac{a}{b} + \frac{\pi - 1}{2}\left(\frac{a - b}{a}\right)^2 - \frac{a - b}{a}\right\}.$$

Since there is only one turn in the circuit this total flux is also the total linkage. For our order of approximation we shall

neglect b wherever it appears subtracted from a and obtain finally

$$L = 4\pi\mu a\left\{\log_\epsilon\frac{a}{b} + \frac{\pi - 3}{2}\right\}$$

when the flux inside the wire is not counted. If we desire to include this inside flux we use the result of § 70 and obtain

$$L = 4\pi\mu a\left\{\log_\epsilon\frac{a}{b} + \frac{\pi}{2} - \frac{5}{4}\right\}.$$

Mutual Inductance of Two Coaxial Circles.

72. An approximate formula for the case of Fig. 44 may be obtained by using the result of § 58; the circles are placed rather far apart and one of them is much smaller than the other. The field at the centre of the small circle due to unit current in the large one is, as we have seen,

$$H = 2\pi y^2/r^3,$$

and as the circle is small we may suppose this expression to hold good with sufficient approximation over the whole area of the smaller circle. Therefore the flux through the one turn of the small circle is

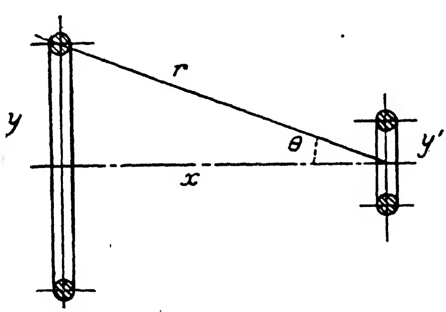


FIG. 44.

$$\pi y'^2 \,.\, 2\pi\mu y^2/r^3$$

or
$$M = 2\pi\mu y'^2 y^2/r^3.$$

Extension to a Circle and Solenoid.

73. The single turn of the small circle may be regarded as an infinitesimal element of a solenoid of radius y' as indicated in Fig. 45. Then if there are τ turns per centimetre the number of turns in a length dx is τdx and the linkages with element dx are reckoned by multiplying this into the flux through a single turn,

already calculated. This contribution to the whole mutual inductance of circle and solenoid may be called dM, and

$$dM = (2\pi\mu y'^2\tau)\sin^3\theta\,dx/y$$
$$= c\sin^3\theta\,dx/y, \text{ say.}$$

Since
$$x = y\cot\theta$$
$$dx = -y\,\mathrm{cosec}^2\,\theta\,d\theta$$
and
$$dM = -c\sin^2\theta\,d\theta$$
$$= -\tfrac{1}{2}c(1-\cos 2\theta)d\theta.$$

Let the length of the solenoid be such that at the far end the

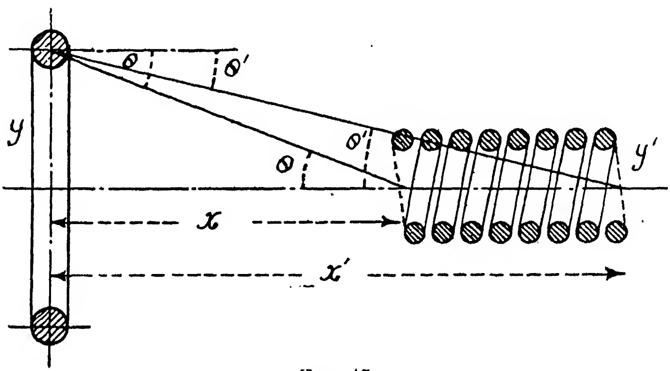


FIG. 45.

angle θ becomes θ'; then we obtain M by integrating from θ to θ'. Thus

$$M = C(\cos\theta' - \cos\theta).$$

This result is useful, for instance, in setting up and using apparatus for the testing of amplifiers.

A further extension of the formula can be made by supposing the circle of radius y to be also an element of a solenoid and treating the value of M last obtained as a contribution to the whole M between two distant coaxial solenoids. But as the integration will take up more space than can be spared it will be left to the reader.

74. The defect of the approximate methods just described is that they do not carry with them any indication of the error that has been made in performing the approximation. Results obtained by such methods must stand under suspicion until in some way or other the limits of the errors have been determined;

and in general they should not be used as a basis for further calculations or experiments without due care.

THE SECOND LAW OF ELECTROMAGNETISM

75. In the preceding section we have learned that electric currents produce magnetic flux in the space around them, and we have used in various ways the conception that the magnetic field possesses energy that might be regarded as the kinetic energy of the moving Faraday lines. This implies that the Faraday lines have inertia, or a property analogous to inertia, and that forces of appropriate type are needed to alter the motion of the Faraday lines. We expect the type of force to be what we have called electric force, of which the line integral is the so-called electromotive force. In this way we have led up to, and indeed have virtually anticipated, the second law of electromagnetism, which is merely a direct and quantitative statement about a very important mode of producing electric force.

The inertia point of view becomes very convenient if we think of the inertia in terms of magnetic flux rather than in terms of the motion of Faraday tubes. To put the matter another way, it is easier for the mind to use the inertia point of view if we picture the inertia of a current in a circuit as due to the spin of its stationary magnetic rings rather than as consisting of a torrent of outward rushing and reflected Faraday lines; and, again, it is mentally easier to deal with the lines of flux of a permanent magnet than with the whirl of Faraday lines required to account for the flux lines. It is easier, for instance, to work out the sizes of the pulleys required for the various machines in a shop by thinking in terms of the revolution speeds of the counter-shafting and mandrels than in terms of the linear velocities of the belting. But it by no means follows that because the one mode is easier it is therefore physically truer than the other.

Using then the notion of spinning flux rings a circuit carrying a current becomes analogous to, say, a freely rotating grindstone, and the opening of a key in the circuit to stop the current resembles the catching hold of the handle. Experience teaches that the handle gives to the hand a blow in the direction of its motion, and analogy suggests that an emf will arise in the electrical case in the same direction as the current, tending to

keep the current running and tending to keep the flux through the circuit at its original value. The energy in the field of the current may, by the process of transference called "work," be converted into an equal quantity of another form of energy, just as the motion of the grindstone might be, and in particular it may become converted into heat just as the energy of the grindstone usually is.

Statement of Second Law.

76. The second law of electromagnetism may be stated thus: Magnetic flux when cutting across an element of length AP produces along AP an electric force equal to the number of lines cutting AP per unit length per second. The permeability and permittivity of the medium, if isotropic, have no effect. The sense of the electric force is shown in Fig. 46. This statement should be compared with that of the first law in § 34, and the figures should be contrasted. By repeating the reasoning of that paragraph and succeeding ones we see that the second law may be turned into other forms. Thus we may say that the emf generated between A and P is equal to the number of lines crossing the element per second without regard to the direction of the lines or the direction of their motion; and by joining elements end to end to form a tortuous curve we may deduce that along any curve, real or imaginary, in a field of moving flux lines the emf at any instant is equal to the number of lines crossing the curve in the same sense per second.

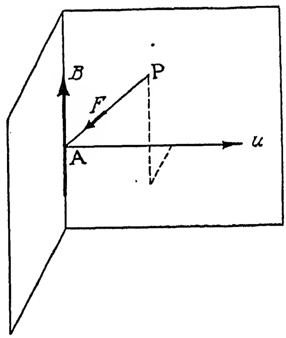


FIG. 46.

Sense of Induced Electromotive Force

77. There are several simple mnemonics for recalling the sense of the induced emf. For example, imagine the eyes of the reader to be the north seeking poles of magnets sending magnetic lines towards the page of this book. On reading the words with the book in the customary position the magnetic lines move from left to right across the page and induce in the paper electric

force in the direction down the page, that is, in the usual direction of reading. Or take a more direct mnemonic based on the suggestion in § 47 that each line of magnetic flux may be con-

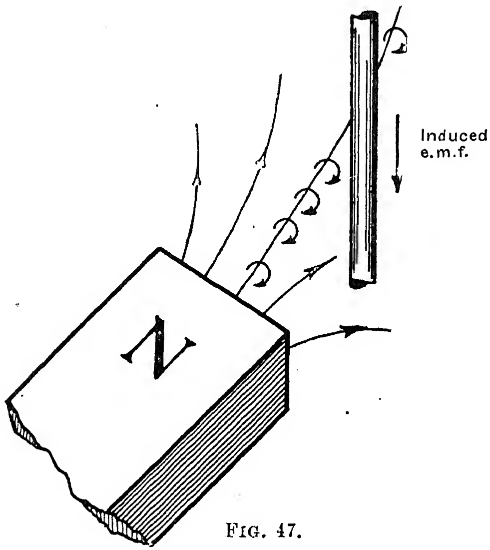


FIG. 47.

ceived as spinning on itself righthandedly. Then, as a line approaches to the curve it is going to cut, the spin gives the clue to the sense of the induced emf. In Fig. 47 the flux line is seen about to cut a wire; the emf induced is the same whether the circuit of the wire is open or closed, or whether the "wire" is of copper or glass. Moreover, the emf is the same in sense when the magnet and its flux lines are stationary and the wire is moved so as to cut them with the same relative motion as before. The first and second laws are contrasted in Fig. 48, wherein is shown

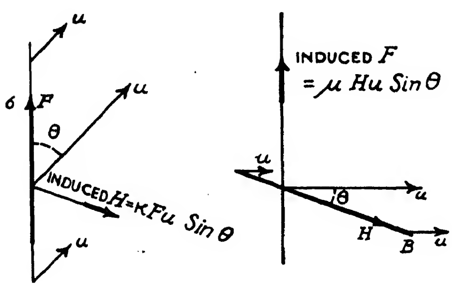


FIG. 48.

the magnetic force produced by a Faraday line moving in any direction, and, quite independently, the electric force produced by a magnetic flux line moving in any direction.

Circuital Form of Second Law.

78. Following the reasoning explained at length in developing the first law we arrive at the circuital form of the second law, which possesses the advantage of escaping the question raised by the idea of relative motion between a magnetic field and a vacuum. It is :—

The electromotive force round a closed curve imagined drawn in any changing magnetic field (or the work done by the electric field on a unit charge guided round a closed curve) is equal to the rate of decrease of the flux linkages with the curve at the instant. The positive direction of the flux lines through a simple curve and the positive direction of the emf along the curve must, in using this law, be taken to be related like the travel and the rotation of a right-handed screw; this is indicated in Fig. 49. It is easy to remember this by use of the remark, in § 75, that the induced emf is in such sense that it tries to generate a current that will keep the total flux through the contour constant. Thus because in Fig. 49 the flux through the circuit is decreasing we imagine positive flux lines pushed through the circuit to make up the loss, and use the ordinary screw rule upon these added lines to show the direction of the induced current. Alternatively we may imagine the flux lines to be moved out of the loop by cutting across the contour, and then we use the rule of Fig. 47.

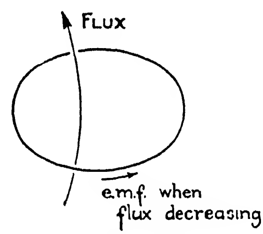


FIG. 49.

When the closed curve is reduced to an infinitesimal plane loop round a particular point in the field and turned so as to obtain the greatest possible emf there at the instant, the emf per unit area is called the "curl" or the "rotation" of the electric force at the point. The second law is then briefly stated as: The curl of the electric force at a point is equal to the rate of decrease of the flux density.

CASES OF INDUCED EMF

79. The circuital form of the second law must be used with care in the case of circuits with several turns, as reference to Fig. 50 will prove. Here lines of flux are indicated threading one and two turns respectively; and plainly, if these are each carried sideways from the circuit, they cut respectively one and two turns and so make different contributions to the emf, according to the earlier way of stating the second law. Evidently in complicated circuits it is the linkages that count, and not merely the flux through the circuit. Thus if the lines are due to a current j in a circuit placed so as to have mutual inductance M with the given circuit, the flux linkages are Mj, and therefore the emf induced in the given circuit is

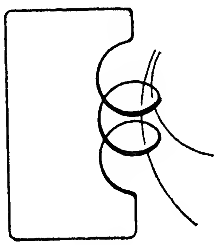


FIG. 50.

$$e = -\frac{d(Mj)}{dt}.$$

Here M may be positive or negative, as explained in § 62. The equation points out that an emf may be induced in a circuit by current in another circuit—

(1) By variation of j alone.

(2) By variation of M alone.

(3) By variation of both M and j.

The first case is that of the stationary transformer—the case which led Faraday to the discovery of the second law; the second is that of the separately excited dynamo (including the unloaded alternator); and the third case arises in many modern forms of dynamoelectric machinery.

It should be added that the flux lines may be due to a permanent magnet instead of to a current in another coil, in which event we have the case of the earliest type of dynamo, the type sometimes called a magneto machine.

80. The law applies not only to the flux linkages made with a coil by external sources of flux, but also to the self-linkages when

current is flowing in the coil. When a current i is running in the coil its self-linkages number Li, as we have seen, and therefore the emf of self-inductance is given by

$$e = -\frac{d(Li)}{dt}.$$

The minus means that if the product Li increases, the induced emf is in the opposite direction to the existing current, and

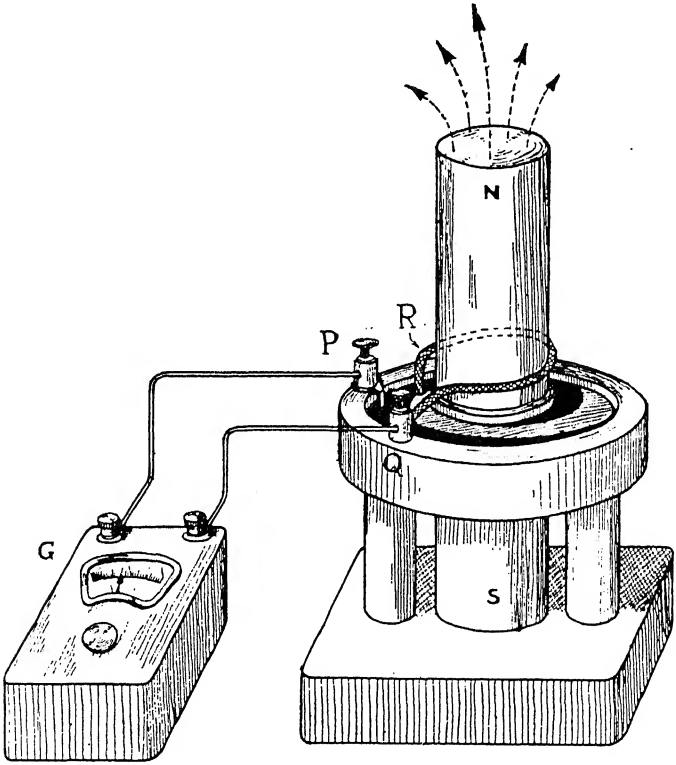


FIG. 51.

vice versâ. This emf may be due, obviously, to variation of L, or of i, or of both. The first case and also the third may be illustrated by the series dynamo, and emfs due to variation of currents in coils occur in innumerable instances. These latter variations are often referred to under the name " counter emf of self-inductance " or " back emf of self-inductance." The back emf has to be overcome by the emf applied to start or stop a

current in a circuit, just as the reaction due to the inertia of a railway wagon, for instance, must be overcome by the force applied to start or stop it. It is of interest to note that in causing a current to grow in a circuit of fixed self-inductance L the work being done at any instant by the applied emf is equal to the back

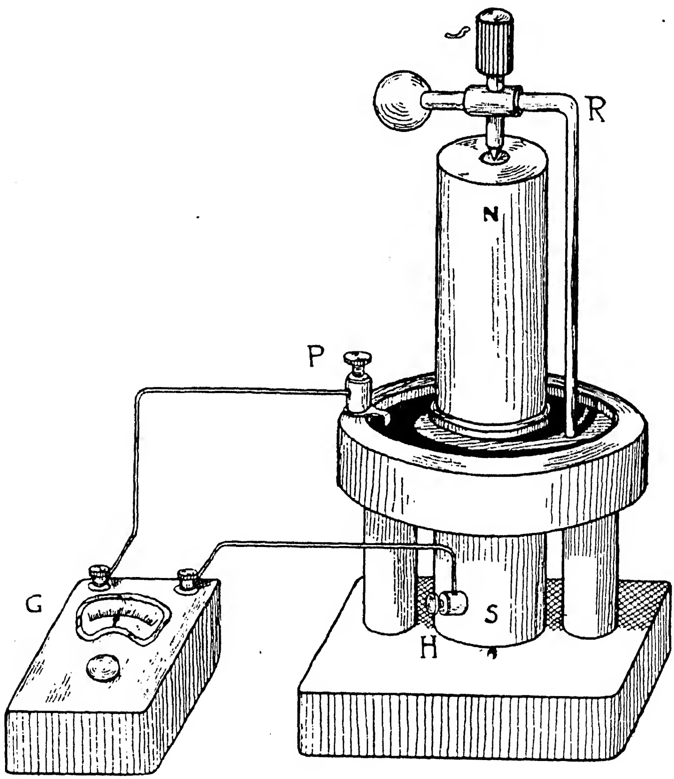


FIG. 52.

emf which it is overcoming multiplied by the current flowing at the instant, that is to say, the power in operation is ei, and

$$ei = Li\frac{di}{dt}.$$

On integrating this from $t = 0$ till any time t we obtain the work done up till the time t. Thus

$$\text{work done by applied emf} = L\int i\frac{di}{dt}dt$$

$$= L\int idi$$

$$= \tfrac{1}{2}Li^2.$$

This, we have seen, is the magnitude of the magnetic energy in the field of the current; all this energy we now see comes from the source of applied emf.

81. Both the above modes of inducing emf arise frequently in our subject—self-linkages and mutual linkages often appear together in the problems discussed in subsequent chapters. Such cases are all included in the equation

$$e = -\frac{d}{dt}\Big(Li + Mj\Big).$$

It is almost unnecessary to mention that the emfs generated by the motion of lines of magnetic flux produce the same effects as the emfs due to other causes, such as voltaic action, thermo-electric processes, etc.; moving lines of flux can therefore produce conduction currents in wires, convection currents of ions, and dielectric currents in insulators. We shall study the last-named phenomenon when we come to electric waves.

82. The circuital theorem as stated above is true only when the flux enters the circuit by cutting its substance, as was emphasised by Faraday. The experiment described by Fig. 51 shows a way of linking a large bundle of flux lines with a circuit without cutting its substance. This circuit PGQR has a flexible portion QR which can be passed round the permanent magnet without cutting the densely packed lines running through the magnet and without breaking the metallic continuity of the circuit. A galvanometer G in the circuit shows no deflection except that due, possibly, to the small current generated by cutting a few stray flux lines. On the other hand, Fig. 52 indicates an experiment in which a movable portion R of a circuit PGHR (completed by the substance of the magnet) is arranged to cut the magnetic flux without altering the linkage with the circuit. The galvanometer deflection is proportional to the speed of rotation of R. These experimental results are made use of in homopolar dynamos.

Induced Currents.

83. The most numerous applications of the second law are those in which the induced emfs act in closed metallic circuits

to produce what may be called induced conduction currents. The induced currents in their turn produce lines of magnetic flux, and completely alter the run of the phenomena. Evidently the amount of complication thus introduced depends greatly upon other circumstances of the circuit, and principally upon the resistance. We shall begin by considering one of the simplest possible cases, namely, that of a very large circle of tubular conductor of which a short portion is shown in Fig. 53. The whole circuit which is completed off the diagram will be taken of resistance R. Suppose that a current is started in this copper tube by means of a battery and that after it reaches full strength the battery is removed without breaking the circuit, as, for example, by short-circuiting the battery. The current will begin to decay, and the magnetic field round it will decay in exact proportion. We can conceive this to happen by the inward movement of the rings of flux—for when the innermost flux ring has collapsed the next ring must follow because of its own tension and because of the unbalanced lateral pressure of the rings about it. Thus the decay of the current is accompanied by an inward motion of the rings of flux. As the rings move they generate electric force according to the second law; and if at any distance x from the axis of the conductor the flux density is B and the velocity inward u, the electric force generated is $F = Bu$. This is true both in the dielectric and in the copper, B and u being, of course, different at different values of x. Symmetry shows that at all places of this nearly straight conductor the emf generated between the points P and Q of Fig. 53 must be the same as that between A and B in order that the equipotential planes in space may be perpendicular to the tube, and therefore we must have, fairly near the conductor,

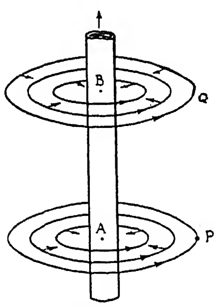


FIG. 53.

$$Bu = F = \text{constant}.$$

But $$B = 2\mu i/x$$

and therefore $$u = Fx/2\mu i.$$

Now F multiplied by the circumference of the circle is equal to

the terminal PD of the resistance R, namely, Ri, and thus we see that u is proportional to x.

84. The value of F at any place such as P or Q in Fig. 53 falls with fall of i and therefore there is a dielectric current between P and Q in the direction from Q to P, that is, in the direction opposite to i. Now the energy in any ring is by § 59 equal to half the total current surrounded by that ring; and the total current, made up of i and all the dielectric current between the ring and the conductor is less than i. We see, therefore, that as a ring shrinks it gathers energy from the dielectric currents it passes through and arrives finally at the conductor with the energy $\frac{1}{2}i$. What becomes of this energy when the flux ring collapses into the copper and disappears? The answer is that when the flux ring cuts into the copper it generates an emf in the same direction as the flowing current, and its energy is dissipated as heat in the portion of the resistance R at the place of collapse. To dissipate all the energy of the field in this way takes time, of course; and since the rate of dissipation is Ri^2 and the stock of energy left in the field is at that instant $\frac{1}{2}Li^2$, where L is the inductance of the circuit, we see that the rate of loss of the stock is proportional to the instantaneous value of the stock and therefore conclude that the decay is logarithmic. To be precise, taking

$$w = \tfrac{1}{2}Li^2$$

$$-\frac{dw}{dt} = Ri^2 = \frac{2Rw}{L}$$

and therefore, by integration,

$$w = w_0 \epsilon^{-\frac{2R}{L}t}$$

where w_0 is the original stock of energy at the time $t = 0$. Since w is proportional to the square of i, we derive

$$i = i_0 \epsilon^{-\frac{R}{L}t}$$

as the law of decay of the current; a law true for any circuit, in fact, whatever its shape.

Flux and Charge.

85. The matter is worth examining in another way. As each flux ring collapses into the copper the emf it generates is an effort to keep the current flowing—an effort made in vain because of

the dissipative effect of the resistance R. Now if along the length of the conductor there are $\mathscr{F}$ rings collapsing per second the emf generated in the conductor is $\mathscr{F}$ units, the contribution to the current is $\mathscr{F}/R$ units, and the quantity of electricity passed along the conductor (in the second) is

$$Q = \mathscr{F}/R.$$

Thus each flux ring as it collapses into the copper squeezes through itself a quantity of electricity of amount $1/R$. The above equation is the basis of the fluxmeter, which is an instrument that measures the flux removed from, or introduced into, a circuit containing the instrument. It is a moving coil galvanometer without a spring control, so the coil stays in the position it is thrown to by the transient current and its displacement is the integral of that current, which is Q. This is all true of any form of circuit.

If the rings shown in Fig. 53 be imagined to have right-handed spin, as already described in § 77, the sense of the spin is seen to be just such as would, so to speak, roll the electricity Q up through them as they sink into the copper. It should be noticed that if R is infinite no electricity is passed, the circuit being interrupted; while if R is zero Q is indefinite for the reason that the current cannot decay in this case, and consequently the rings never collapse.

Reaction of Induced Currents.

86. It has been remarked before that the instances in which the induced emf produces current are very numerous in electrical practice. The flow of current introduces a complication not yet described. In Fig. 54 a pair of circuits is pictured, one of which, called the primary, contains a voltaic cell and a switch which is short-circuited by a high resistance, and the other, called the secondary, is in the plane of the primary and may be closed or opened at will. Such a pair of circuits is said to be coupled and is spoken of as a transformer, when constructed for actual service. Let the primary key be closed and the current be steady and the magnetic field be fully formed as indicated in the figure. Let the secondary key be open. Now let the primary key be opened suddenly. The current will be practically stopped by the high resistance now thrown into the circuit. All the flux rings will

collapse into the primary wire, generate a transient emf and current in the primary circuit and give their energy to heating the resistance; and the rings that thread the secondary will, in collapsing, cut that circuit and generate an emf in it. There will be no secondary current because the circuit is open. It is permissible to assume that the presence of the secondary does not affect the collapsing of the flux rings to a significant extent and that therefore the average emf generated in the primary is unaffected by the open secondary.

87. But now let the secondary key be closed and the experiment repeated. Noting that the induced secondary current tends to

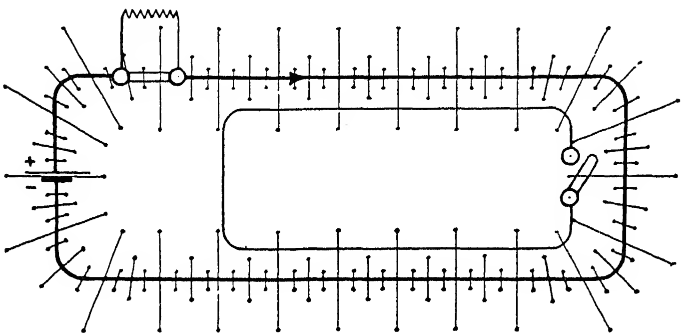


FIG. 54.

keep constant the flux through that circuit, we see that it will run in the same sense as the primary current did; and the flux rings due to the secondary current will, as they expand from the secondary wires as centres, cut the primary wires. But these new rings, where they cut the primary, are directed oppositely to the collapsing rings of dying primary current, whence we conclude that the emf now generated at any instant is less than at the corresponding instant in the preceding experiment with the secondary circuit broken. Moreover, if the secondary circuit possesses resistance, the joulean wastage in it must be derived from the energy of the primary flux rings trapped in the circuit, and so much less reaches the primary resistance. All this influence exerted by the secondary on the primary circuit is called the reaction of the secondary on the primary. The presence of the secondary reduces the emf, the current, and the heating in the

primary, and this reaction can be summed up by saying that the secondary reduces the effective inductance of the primary.

88. It is not difficult to work out the progress of the currents and voltages throughout the experiment, but instead of doing this consider the extreme case when the secondary circuit is a perfect conductor. First suppose that the flux is threading the secondary circuit as in the figure, that this circuit is closed and the primary opened. Then all the lines not threading the secondary collapse into the primary, but the others cannot escape from the secondary; for if any minute fraction of the whole of this trapped flux be supposed to cut the perfect conductor it will generate a corresponding emf and this will produce a current just big enough to make up for the loss, and this current will never decay: therefore no fraction of the flux escapes. In the same way if the perfectly conducting circuit were given closed and were submitted to the action of the primary with the intention of causing primary flux to thread it, we should find it impenetrable. The only way of linking flux with this perfectly conducting secondary would be to leave its key open so that the flux rings from the primary could leak in. Passing from this extreme case it is evident that if the secondary possesses some slight resistance the flux rings will in fact gradually all collapse, some into the primary, some into the secondary, and it is clear that the process will be more prolonged the better the secondary conductance. This has had direct application in various delay-action relays. For instance, thick copper tubes round the legs of a horseshoe electromagnet of which the "keeper" is the moving tongue of the relay have the effect of delaying by several seconds the instant at which the tongue moves, whether from the condition of full current or no current through the windings; for in either case large currents arise in the low resistance tubular circuits and delay the change imposed by breaking or making the current through the windings.

Steady Induced Current.

89. It is of interest and of importance at a later stage of our subject to picture the interaction between the inducing flux and the flux belonging to the induced current. The simplest way of studying the matter is by the relative motion of the perfectly conducting wire and the magnetic field of Figs. 55 and 56. Let

us suppose the wire to be moving from right to left; then the induced current is through the paper away from the reader, as indicated by the screw head, and its magnitude is determined by the rate at which the moving conductor cuts the flux lines and by the resistance of the external circuit. Fig. 55 shows the flux lines of the magnet and the flux rings of the current separately; Fig. 56 represents these fluxes compounded. At the instant shown the innermost ring of flux has been split off from A, the next from B, and another ring is in process of formation. These complete rings are evidence that a current is flowing. Each ring is torn from the field of the magnet by means of the lateral pressure exerted by the rings already formed.

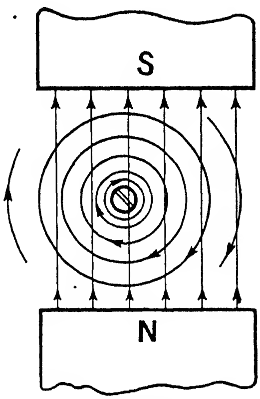


FIG. 55.

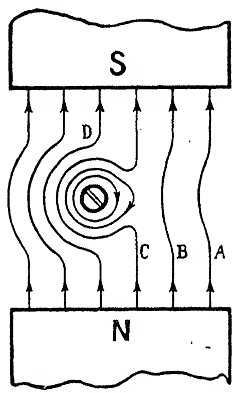


FIG. 56.

90. If the wire in Fig. 55 is moved so as to cut a constant number of flux lines per second, a constant emf E is induced, and a steady state is soon set up as regards the current I and its flux rings. The flux pattern of Fig. 56 will be a moving one and as each new outer ring is acquired the inner rings will contract and the innermost one will collapse into the wire. Thus the induced emf E is equal to the number of new rings acquired per second. But the power being dissipated in the external resistance is known to be EI, from which it follows that each new ring must bring to the circuit I units of energy. This contribution of I ergs per ring (in a cgs system of units) should be contrasted with the result of § 59, where $\frac{1}{2}I$ appears because the average work done in building the magnetic field by, say, a voltaic cell is there con-

sidered, and some rings are therefore formed when the current is very small, some when it is near the final value, and others at intermediate stages. In the present case each ring is linked up at the full strength of the current.

91. The energy left in each flux ring and finally delivered to the external resistance and dissipated as heat is supplied by the agency moving the wire. Inspection of the curved lines of force in Fig. 56 indicates that the motion of the wire is resisted by the lateral pressures between the flux lines on the left and by the longitudinal tensions in the lines such as those near C and D. Let Y represent the total resultant mechanical force on unit length of the wire (plainly symmetrical and to the right) and also the mechanical force per unit length needed to move the wire to the left; then the rate of working is Ylu, where u is the broadside velocity of the wire and l its length. This must equal the energy being steadily delivered to the external resistance. Therefore

$$Ylu = EI.$$

But the induced emf E is given by

$$E = Blu$$

where B represents the flux density.

Thus $$Y = BI.$$

Electrodynamic Force

92. The flux rings and lines of Fig. 55 may be produced by holding the wire quite stationary between the magnet poles and sending a current through it from a voltaic cell. The resultant flux curves of Fig. 56 and the mechanical forces exerted on the wire are therefore the same when the current I is due to a voltaic cell as when it was due to relative motion. The mechanical force exerted by a magnetic field on a conductor conveying a current is called electrodynamic force. In general its magnitude at any point of a circuit is equal per unit length of the conductor to the current multiplied by the component of flux density there perpendicular to the current, as indicated by the formula given above. It is perpendicular to both current and flux, and its sense is settled by observing that the flux lines of the field and of the current repel each other laterally when running in the same general direction, and *vice versâ*. In this way it is easily seen from a

sketch that two parallel wires conveying currents in the same direction are attracted, while when conveying oppositely directed currents they are repelled.

93. It is an immediate consequence of the above reasoning that mechanical work is done by the electrodynamic forces when the wire is allowed to move. The apparatus becomes in fact an electromotor. The only possible origin of the energy for this work is the source of emf that keeps the current constant. This is brought into play by the operation of the second law; for when the wire moves it cuts flux lines and therefore an emf is induced in it, and the rules already given, when applied to Fig. 55, show that this emf will be in the opposite direction to the current; hence in order to keep the current constant the emf of the source must increase to counter the new induced emf, and the power thus called out is equal at each instant to this emf multiplied by the constant current. This statement may be written :—

Electrodynamic power = current × rate of cutting of lines.

94. In a complicated circuit these events occur at every centimetre of its windings. We know already from § 79 that the total rate of cutting lines of flux by any closed circuit is equal to the rate of increase of the flux linkages between the circuit and the field, and therefore we see that the result of the last paragraph can be written for any instant of time :—

Power expended by means of the electrodynamic forces on the moving circuit = the current in it × the rate of change of the linkages of circuit and field.

95. Very familiar instances of this principle are seen in the ordinary direct current motor and in the moving coil galvanometer. In both these instances a coil, movable about a symmetrical axis in the plane of the coil, is placed in a magnetic field perpendicular to the axis. Let Fig. 57 represent a view along the axis, the coil being taken of only two turns for simplicity and placed in any general position in the magnetic field. Let the current be led through the coil (say by aid of leads along the axial support, not shown) in the direction given by the slots and dots. The screw rule tells that magnetic lines in the clockwise direction will arise around the upper wires in the figure, and that lateral repulsions between these and the field lines above will produce a torque in the sense indicated by the curved arrows. The value of the torque depends on the angle between the coil

and the field. If the coil is free it will turn till the torque is zero. This final position is shown in Fig. 58 where the electrodynamic forces on the coil are in equilibrium. In this position the flux due to the current in the coil threads the coil in the same direction as the field does, that is, the field is linking with the coil in the positive direction as defined by the current. The behaviour of the coil can therefore be described in the statement:—

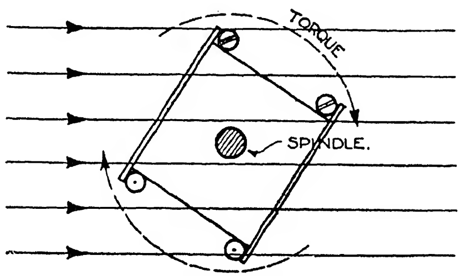


FIG. 57.

A coil carrying a current in a magnetic field tends to place itself so that as much flux of the field as possible is linked with the coil in the positive direction.

This is often expressed by saying that a coil carrying a current in a magnetic field behaves like a magnet, one side of the coil being of north-seeking polarity, the other of south-seeking.

95. By integrating throughout the whole time of motion the statement at the end of § 94 we see also that:—

The work done by the electrodynamic forces is equal to the constant current in the coil multiplied by the total change of its flux linkage with the field.

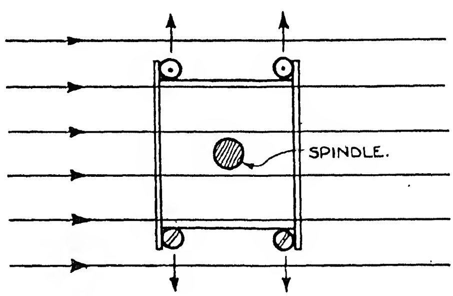


FIG. 58.

It may be well to remark at this point that the work done by the source of emf in setting up its own magnetic field is always in accordance with § 61, equal to $\frac{1}{2}Li^2$ whether there is an external magnetic field or not; it is only when the coil moves that the extra call is made upon the source.

Extension to any Circuit.

97. The above conclusions can be extended by precisely the same reasoning to the study of flexible circuits or circuits with movable parts placed in their own magnetic field ; the circuit will deform, and its movable parts will move, so that the circuit increases its linkage with its own magnetic field. For example, the bridge of copper floating in the mercury channels in Fig. 59 will move away from the remainder of the circuit when current is flowing. Of course this is nothing more than a manifestation of the mutual lateral repulsions of the flux rings that pack them-

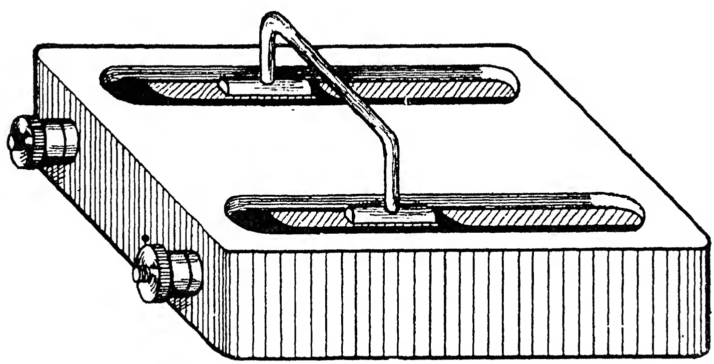


FIG. 59.

selves into the circuit when the current flows round it. We may illustrate this view by the electromagnet and the armature of a series motor, which constitute a circuit that is encouraged to enlarge its self-linkages indefinitely by the action of the commutator.

Electrodynamic Forces on Induced Currents. Lenz's Rule.

98. There are numerous cases in which the current in the moving conductor is not kept constant by an independent source, and many in which all the current is being induced by the motion of the coil relative to the field. Let us examine the latter case by imagining a closed coil of wire or a piece of metal to be pushed by hand into a fixed magnetic field. The currents induced will do work in heating the conductors, and clearly the energy for

this must come from the hand. In fact the electrodynamic forces arising between the field and the induced current will offer to the hand a force tending to stop the motion. This deduction is often called Lenz's rule for the direction of the induced current or the direction of its electrodynamic forces. As an instance of this we may take the armature of any dynamo, which, the rule shows, has its rotation opposed by electrodynamic forces when current flows.

UNITS

99. The electromagnetic formulæ of the preceding pages hold good for any consistent system of units based on the units of length, mass and time as fundamental units. The two principal theoretical systems of units are called the electrostatic and the electromagnetic. In the electrostatic system the starting point is the unit of electricity, this unit being such that two of them concentrated at points 1 cm apart in a vacuum repel each other with a force of 1 dyne. Reference to § 3, II., shows that this definition implies that the inductivity (permittivity) of vacuum is taken equal to unity in this system. In the electromagnetic system of units the starting point is the unit of a fictitious magnetism, and the unit is such that two of them concentrated at points 1 cm apart in a vacuum repel each other with a force of 1 dyne. Reference to § 30, II., shows that this definition implies that the permeability of vacuum is taken equal to unity in the electromagnetic system. By aid of the two laws of electromagnetism all the electric and magnetic quantities may have their units of measurement settled on either the electromagnetic or the electrostatic basis; and when this is done for any particular quantity it is found that the electrostatic and the electromagnetic units of that quantity are very different in absolute magnitude. For instance, it is found that two magnetic poles, each measuring one unit in the electrostatic system, placed 1 cm apart in vacuum do not repel each other with a force of 1 dyne, but, in fact, with a force of 9×10^{20} dynes. Each is therefore equal to 3×10^{10} electromagnetic units of magnetism. Again, two separate units of electricity in the electromagnetic system when placed 1 cm apart in vacuum repel each other with a force of 9×10^{20} dynes; and there-

fore a unit of electricity in the electromagnetic system is 3×10^{10} times as great as a unit of electricity in the electrostatic system. These statements have been confirmed by fairly direct experiment; and many observers have determined the ratios of the units of other quantities in these two systems. We proceed to apply to this problem certain theoretical considerations.

100. Let us take the first law and apply it to the Faraday lines of an electric field F moving at right angles to itself with velocity u. It produces a magnetic field of intensity H which is perpendicular to F and u and can be represented by lines of flux of density μH per square centimetre. Let us assume that these lines of induction accompany the Faraday lines; then their motion must, in accordance with the second law, give rise to electric force of amount μHu in the direction of F. This electric field is not, however, superposed upon the original field; it must be regarded as a part already included in F. The question at once arises: At what speed must an electric field move in order that the magnetic field it generates may in turn by its equal motion account for the whole of the electric force? The question is answered by writing

$$F = \mu Hu = \mu(\kappa Fu)u$$

or

$$\mu\kappa u^2 = 1.$$

Hence at the speed $(\mu\kappa)^{-\frac{1}{2}}$ it is immaterial whether we look upon the system as a set of moving Faraday lines or as a set of moving lines of magnetic flux: in the one case we regard the moving Faraday lines as causing and supporting the magnetic field, and in the other we regard the moving induction as causing and supporting the electric field.

101. At this special value of u, referred to in § 48, and to which we shall assign the symbol s, the electric and magnetic fields can travel without free electricity or magnetism. The reasoning of § 49 shows also that the mechanical effects of the electric and magnetic lines are cancelled except in the direction of motion. This unsupported movement of the states called electric and magnetic force and strain or flux is an instance of the electromagnetic wave motion first discussed quantitatively by Maxwell. We have just seen that the speed of these electromagnetic waves is

$$s = \frac{1}{\sqrt{\mu\kappa}}.$$

This speed has been measured by actual experiments on electric waves in the laboratory and also by wireless telegraphy between Arlington and the Eiffel Tower at Paris. It is found to be very nearly 3×10^{10} centimetres per second in air, which is also the speed of light. The equation above enables the velocity to be calculated in any other medium when the constants μ and κ are known.

Inductivity and Permeability in Various Systems.

102. The above equation enables us to compare the magnitudes of the constants μ and κ of a given medium in various systems of units, or, what comes to the same thing, to find the relative sizes of the units of permittivity and permeability. Let us take vacuum, or air, as the medium.

In the electrostatic system of units κ is taken arbitrarily equal to 1 and therefore since $s = 3 \times 10^{10}$ centimetres per second,

$$\mu = \frac{1}{\kappa s^2} = \frac{1}{9} \times 10^{-20}.$$

In the electromagnetic system of units μ is taken arbitrarily equal to 1 in a vacuum, and therefore the unit of permeability in the electrostatic system must be very large to yield so small a number as that just arrived at for the permeability of vacuum; in fact the electrostatic unit of permeability is 9×10^{20} times greater than that used in the electromagnetic system.

103. In future we shall consider s to be the velocity of electric waves in vacuum or air so that

$$s = 3 \times 10^{10}.$$

Hence the ratio of the units of permeability is s^2, the electrostatic unit being the larger.

104. On the other hand, in the electromagnetic system the unit of permittivity is large compared with that in the electrostatic system. For in the electromagnetic system, where $\mu = 1$, we have

$$\kappa = \frac{1}{\mu s^2} = \frac{1}{s^2}$$

for the permittivity of vacuum. This small numerical value

shows that the electromagnetic unit of permittivity is s^2 times as great as the electrostatic unit.

105. In the " practical " system of units arbitrary multiples and submultiples of the electromagnetic units have been taken to form a system. For instance a farad, the practical unit of capacitance, is equal to a thousand-millionth part of an electromagnetic unit of capacitance. Now the capacitance of a condenser is determined solely by its dimensions and the permittivity of its dielectric, and in all three systems the dimensions are measured in centimetres ; hence if a vacuous condenser of unit capacitance in the electromagnetic system must be called 10^9 farads the permittivity of vacuum in this practical system of units must be $10^9/s^2$ or $\frac{1}{9} \times 10^{-11}$. This shows that the practical units of capacitance and inductivity are 9×10^{11} times larger than the electrostatic units. The speed equation now gives the permeability of vacuum as

$$\mu = \frac{1}{s^2(10^9/s^2)} = 10^{-9},$$

proving that the practical unit of permeability is 10^9 times greater than the electromagnetic unit.

Conversion from One System to Another

106. In wireless telegraphy calculations, more than in any other branch of applied electricity, it is necessary to be able to use all three systems of units and to move from one to the other. For instance the air or glass condensers used in wireless telegraphy are conveniently measured in electrostatic units (called the centimetre), but are often rendered in billifarads ; the coils in a radiotelegraphic station may be of magnitude one or ten thousand electromagnetic units (called the centimetre), but are often given in millihenrys. Resistances, however, are nearly always stated in terms of the practical unit called the ohm. Indeed it is quite frequent to have the data of a set of wireless telegraph apparatus given with the inductance and permeability in electromagnetic units, capacitance and inductivity in electrostatic units, resistance and reactance in practical units. It is therefore necessary to prepare for rapid transference from one system of units to the other.

Suggested Names of Units.

107. One of the difficulties in explaining and using the connections between the various systems of units arises from the absence of names for many of the units. Names have, however, been given to the principal practical units, for instance the unit of capacitance is called the farad. We have already seen that one electromagnetic unit of capacitance equals 10^9 farads and therefore we might call the electromagnetic unit of capacitance "one 10^9 farad." The electrostatic unit of capacitance is $(10^9/s^2)$ farad and might be so called in speech. This nomenclature would, perhaps, be too great a strain on the memory; so it would be easier to call the electromagnetic unit of capacitance "one emfarad" and the electrostatic unit "one esfarad"; and then the conversion factors could be obtained as required from a table, such as that below.

Calculation of Conversion Factors.

108. The mode of finding the conversion factors must be explained before giving the table. The most useful rule to follow is to take the quantity for which the conversion factor is required and try to express it in terms of purely mechanical quantities and (or) electrical quantities already known. This method was in fact applied in § 105, for there it is shown that the capacitance of a definite condenser divided by the electric inductivity must be a certain length, and therefore in all systems in which the unit of length is the same this quotient must have the same numerical value; and if the unit of length is changed the change in the numerical value of the quotient is easily allowed for and the ratio of the units determined.

Let it be required to connect the unit of electric field strength F, which is the same physical entity as potential gradient, in the three systems. We shall utilise the fact that the energy stored per unit volume in any electric field is $\kappa F^2/8\pi$. In the two cgs systems this is in ergs, but in the practical system it is in joules. For a particular electric field in air the numerical value of the above expression must be, of course, the same number of ergs in all systems. For instance if in this field the potential gradient happens to be 1 volt per centi-

metre, which is one practical unit of electric force, the above expression must have substituted in it $F = 1$ and $\kappa = \frac{1}{9} \times 10^{-11}$, giving

$$\frac{10^{-11}}{8\pi \cdot 9} \text{ joule or } \frac{10^{-4}}{8\pi \cdot 9} \text{ erg.}$$

But in the electrostatic system the same formula gives the result in ergs direct, and in this system $\kappa = 1$. Hence

$$\frac{\kappa F^2}{8\pi} = \frac{F^2}{8\pi} = \frac{10^{-4}}{8\pi \cdot 9}$$

from above, or $$F^2 = \tfrac{1}{9} \times 10^{-4}$$

or $$F = \tfrac{1}{300}.$$

Verbally, one practical unit of electric force equals 1/300 electrostatic unit. We conclude, also, that one electrostatic unit equals 300 volts per centimetre. We have

$$1 \text{ esvolt per cm} = 300 \text{ volts per cm}$$

whence, also,

$$1 \text{ esvolt} = 300 \text{ volts},$$

which gives the conversion factor for electric potential difference.

Turning now to the electromagnetic system, since the inductivity of air is $\frac{1}{9} \times 10^{-20}$ in this system we have

$$\frac{\kappa F^2}{8\pi} = \frac{10^{-20} F^2}{9 \cdot 8\pi} = \frac{10^{-4}}{8\pi \cdot 9} \text{ erg},$$

from above. Therefore $$F^2 = 10^{16}$$

and $$F = 10^8.$$

Hence

$$1 \text{ emvolt per cm} = 10^{-8} \text{ volt per cm}$$

and $$1 \text{ emvolt} = 10^{-8} \text{ volt}.$$

We have thus obtained the ratios of the units of electric force and electric potential difference in all three systems.

109. From the units of field strength we may pass to those of electric strain density; from potential difference to charge, since $Q = CV$; current must have the same conversion factors as charge because current is of the nature of charge per second in all the

systems. Moreover, magnetic field strength must have the same conversion factors as current because the field due to a current in a wire is a numerical, or, rather, angular multiple of current divided by a distance in centimetres. The factors for resistance follow from these obtained for voltage and current, for flux from permeability and magnetic field, and inductance from flux. In this way the following table has been constructed.

Connections between Various Units.

110. In the table the names of the units are given and certain of their connections indicated. But there are other connections; for example, the joule is the watt-second, an ampere is a coulomb per second. The reasoning given fully above shows that the practical unit of electric inductivity is that of a medium in which a sphere of 1 cm radius would have a capacitance of 1 farad, and therefore the unit of electric inductivity is a farad per centimetre. In the same way § 109 leads to the conclusion that the practical unit of permeability is a henry per centimetre. Other connections are supplied by the equations of preceding sections. They are collected here for reference.

Practical unit of electric inductivity (permittivity)	= farad/cm.
Practical unit of permeability	= henry/cm.
Henry × ampere	= volt × second.
∴ Henry	= second × ohm.
Farad × volt	= ampere × second.
∴ Second	= farad × ohm.
Joule	= watt × second.
Coulomb	= ampere × second.
,,	= farad × volt.
Volt	= ohm × ampere.

111. Conversion Factors for the Practical, the Electromagnetic and the Electrostatic Systems of Units.

	Practical Unit and Symbol.	Electro-magnetic Unit in terms of practical.	Electrostatic Unit in terms of practical.	
Capacitance .	Farad, F .	10^{9} practical units	$\frac{1}{9}\times 10^{-11}$ practical unit	1 esfarad = 1 cm. 1 jar = 1000 cm.
Conductance .	Mho . .	„ „	„ „	
Conductivity .	Mho/cm³ .	„ „	„ „	
Inductivity or Permittivity	Farad/cm .	„ „	„ „	
Electric current	Ampere, A .	10 „	$\frac{1}{3}\times 10^{-9}$ „	
„ quantity	Coulomb, C .	„ „	„ „	
„ strain density	Coulomb/cm² .	„ „	„ „	
Magnetic force	Decigauss / cm or ampere/cm	„ „	„ „	
Magnetomotive force . .	Decigauss or ampere-radian	„ „	„ „	
Energy . .	Joule, J .	10^{-7} „	10^{7} „	1 erg = 10^{-7} joule.
Mass . .		„ „	„ „	
Power . .	Watt, W .	„ „	„ „	
Electric force .	Volt/cm .	10^{-8} „	3×10^{2} „	
Electromotive force or voltage	Volt, V .	„ „	„ „	
Flux . .	Weber . .	„ „	„ „	
Flux density .	Maxwell .	„ „	„ „	
Magnetic pole.	Volt-second (per radian)	„ „	„ „	
Inductance .	Henry, H .	10^{-9} „	9×10^{11} „	1 emhenry = 1 cm.
Permeability .	Henry/cm .	„ „	„ „	
Resistance, Reactance and Impedance .	Ohm, Ω .	„ „	„ „	

Pico, p $= 10^{-12}$, billi, b $= 10^{-9}$, micro, $\mu = 10^{-6}$, milli, m $= 10^{-3}$, kilo, k $= 10^{3}$, mega, M $= 10^{6}$.

Mode of using Conversion Factors

112. The easiest way of illustrating the use of the table is by converting 47 miles into yards. To do this we may write

$$x \text{ yards} = 47 \text{ miles},$$

where x has to be determined. We have

$$x = 47 \frac{\text{miles}}{\text{yards}}$$
$$= 47 \frac{1760 \text{ yards}}{\text{yards}}$$
$$= 47 \times 1760$$
$$= 82720.$$

As an example, find the value of

$$\frac{C_1 C_2 R^2}{4L}$$

when $C_1 = 2$ jars, $C_2 = 1{\cdot}5\ \mu F$
$R = 12\Omega$, $L = 18{,}000$ cm.

The expression is equal to

$$\frac{2 \times 1{\cdot}5 \times 12^2}{4 \times 18 \times 10^3} \times \frac{\text{jar} \times \mu\text{F} \times \text{ohm}^2}{\text{cm}}$$

or
$$\frac{6}{1000} \times \frac{1000 \times \frac{1}{9} \times 10^{-11}\text{F} \times 10^{-6}\text{F} \times \Omega^2}{10^{-9}\text{H}}$$

or
$$\frac{6}{1000} \times \frac{10^{-5}\text{F}^2\Omega^2}{9\text{H}}$$

or
$$\tfrac{2}{3} \times 10^{-8} \frac{\text{F}^2\Omega^2}{\text{H}},$$

where F stands for farad, Ω for ohm, and H for henry. But by § 110, FΩ = second, and H = second $\times$ ohm, and therefore

$$\frac{\text{F}^2\Omega^2}{\text{H}} = \text{F}.$$

The answer to the question is thus

$$\tfrac{2}{3} \times 10^{-8} \text{ farad}.$$

CHAPTER III

THEORY OF ALTERNATING CURRENT AND OSCILLATIONS

Definitions. Crank Diagrams.

1. An alternating current is one that flows first in one direction and then in the other. Usually it is understood that all the values, positive and negative, are repeated continually at equal intervals of time. The time of a complete repetition is called the " periodic time " or " period," and is given in this book the symbol T. When the periodic time is less than about a ten-thousandth of a second the alternation may be called a continuous

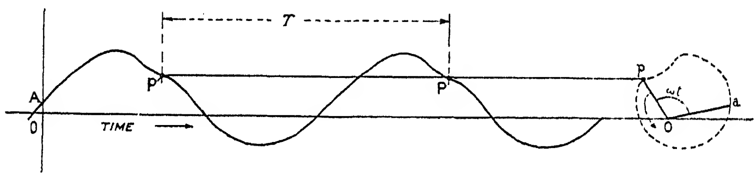


FIG. 60.

oscillation. Fig. 60 shows the graph of an alternating current plotted on a time base with two points P′, P marked on it at corresponding points in two consecutive repetitions. The time between any two such points is a period. There is another way of representing a periodic curve, and this is indicated at the right-hand side of the same Figure. Here a revolving crank Op, of variable length, is supposed to be following the dotted outline while rotating with constant angular velocity, and the dotted outline is so drawn that the height of the end p above the datum line gives an ordinate in the periodic curve, while the time that has elapsed since the crank left its initial position gives an abscissa. This initial position, sometimes called the epoch position, is indicated at Oa, and corresponds to the origin of time O of the curve P′P. The crank has been round about $1\frac{1}{2}$ times while the curve has been described from O up to P. In electrical theory

a complete period is often called a cycle, and the number of cycles per second is called the frequency (symbol f). If ω represents the angular velocity, that is, the angle described per second, then since there are 2π radians in one revolution the angular velocity is given by

$$\omega = 2\pi f.$$

The angular velocity ω is sometimes called the frequency per 2π seconds. Also since f is the number of cycles per second, the time of a cycle or the period is $1/f$. Thus

$$f = 1/T = \omega/2\pi.$$

2. Representing the alternating magnitude by aid of a variable crank suggests very clearly the simplest possible kind of alternation. It is that represented by the uniform rotation of a crank not varying in length. This is indicated on the right hand of Fig. 61, where the crank Op has its end p describing a circle.

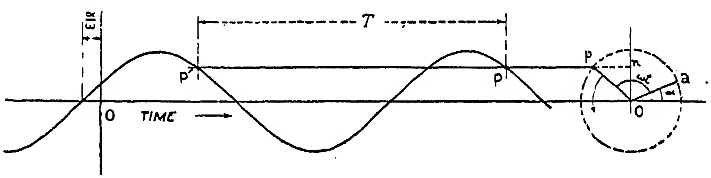


FIG. 61.

In this Figure it is supposed that the plotting of the corresponding periodic curve started when the crank was in the position Oa. Since the height of the point p above the datum line, or, in other words, the projection On on a perpendicular to the datum line, is equal to the sine of the angle $\omega t + \alpha$, the curve is called a sine curve. The angle marked α appears on the left-hand side of the origin as a time of magnitude α/ω. A current or voltage whose values follow the variation indicated by the sine curve is said to be a simple harmonic alternation or oscillation. Let i represent the current at any instant indicated by t, and I represent the length of the crank; then,

$$i = I \sin(\omega t + \alpha).$$

Here I is called the "amplitude," or the "maximum value," or the "crest value," and α the "phase angle." It will be noticed that these quantities are given just as plainly by the crank diagram as by the curve; in fact, anything that can be deduced from the curve can be deduced from the crank diagram.

On comparing the equation with the curve it will be seen

that, in order to obtain the zero value from which the sine curve may be conveniently said to start, it is only necessary to equate the angle $\omega t + \alpha$ to zero (which implies zero value for the sine), and thus obtain $t = -\alpha/\omega$ as the abscissa of the starting point.

Harmonics.

3. Simple harmonic curves can be combined to make complicated curves if the constituent curves are not all of the same period. When they are of the same period another simple harmonic curve arises, as will be seen later. The most important cases of combination of sine curves of different frequency arise when the frequencies of the constituent curves are all whole number multiples of the slowest one. In such a combination the slowest curve is called the fundamental. In music these constituents of multiple frequency are harmonious with the fundamental and are therefore called harmonics. The harmonic of double frequency (if it exists) is called properly the first harmonic, and also the double harmonic. The constituent of triple frequency is called the triple harmonic, and so on. The fundamental is often called in acoustics the first partial, the double harmonic the second partial, and so forth. Fig. 62 shows the crank method of picturing a curve possessing a fundamental and two harmonic constituents. On the first crank Op a second crank pq is added which rotates with respect to the first just as often as the first rotates with respect to the datum line. The third crank qr rotates also at the same rate with regard to its predecessor. Thus the angular velocities of the fundamental and its two harmonics are ω, 2ω, and 3ω. The result is equivalent to the rotation of a crank Or, and the projection of this crank gives on a time base a curve which must be periodic, because after a time equal to any whole multiple of the complete period of the fundamental all three cranks will be in the same position again. Fourier showed that any magnitude that was periodic

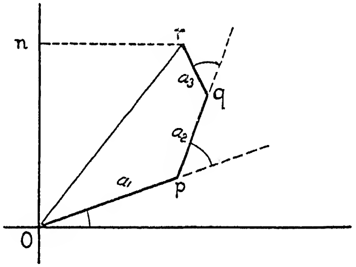


FIG. 62.

and single valued, and that varied without being interrupted or becoming infinite, can be represented uniquely by a fundamental sine curve with properly proportioned harmonics. He gave a mathematical process for dissecting any periodic function, and this process is now called Fourier analysis. It is described and applied in § 33 *et seq.*

Phase of Sines of Equal Period.

4. Simple harmonic curves of the same period can differ in amplitude and in phase. By the term " phase at a given instant,"

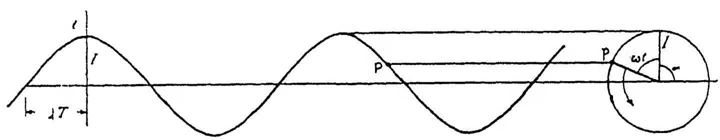


FIG. 63.

when interpreted quantitatively, is usually meant the fraction of the whole period that has elapsed since the curve last passed from a negative to a positive value through zero. Phase can be reckoned in angle as well as in time, and this is often convenient in the crank diagram. The " phase difference " of two simple harmonic quantities, which can only have a definite meaning

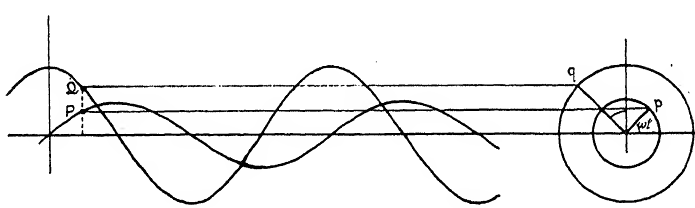


FIG. 64.

when the alternations are of the same frequency, is the difference of their phases reckoned at the same instant, or, in other words, it is the amount reckoned in time or angle by which one quantity anticipates the other in passing through the nearest corresponding maximum, minimum or other value. It is worthy of mention that any cosine curve may be regarded as a sine curve with its phase increased $\frac{1}{2}\pi$ radians or, reckoned in time, $\frac{1}{4}T$, the cosine maxima, etc., occurring earlier than the corresponding sine maxima. This is indicated by Fig. 63. Two curves

differing in phase by a quarter period, as indicated in Fig. 64, are said to be in quadrature. In the corresponding crank diagram the marked angles are equal. It should be noticed also that two alternations having all their simultaneous values of opposite sign

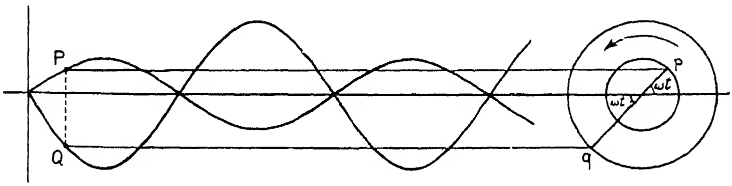


FIG. 65.

are said to be " opposite in phase " ; their phase difference is π in angle or $\frac{1}{2}T$ in time. This is illustrated in Fig. 65.

Sum of Two Equi-frequent Sines.

5. If two sine curves having equations :—

$$i_1 = I_1 \sin \omega t,\ i_2 = I_2 \sin (\omega t + \phi)$$

be plotted on a time base and the ordinates added, the result will be another sine curve. Instead of carrying out the plotting we may use a crank diagram as indicated in Fig. 66. The addition may be effected by drawing the parallelogram and its diagonal OI, or merely, as shown alternatively in the same Figure, by adding I_2 on the end of I_1, as is done in mechanics in using the triangle of forces. Electricians have become accustomed to call this " vector addition " and to call the cranks " vectors." It should be noticed that Fig. 66 differs from Fig. 62

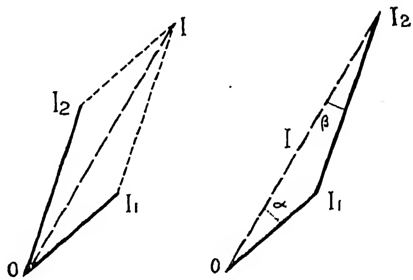


FIG. 66.

by the fact that the angles marked in Fig. 66 are invariable, while the marked angles in Fig. 62 increase with time because the successive vectors rotate at different speeds. It is clear that OI obtained by either of the equivalent methods is constant and rotates at the same speed as either of the constituent vectors,

that is, at angular velocity ω. The usual projection method will give a sine curve of greater amplitude than either constituent and of intermediate phase. The angle α indicates the phase difference of the resultant and one constituent, and since OI will become perpendicular to the datum line before OI_1 does, we say that the former leads the latter. On the other hand, the constituent OI_2 is seen from the left-hand diagram to become perpendicular to the datum line before OI does, and therefore we say that OI lags behind OI_2. The same results may be arrived at algebraically by the well-known formulæ for the sum of two sines. Starting from the above equations we have :—

$$i_1 + i_2 = I_1 \sin \omega t + I_2 \cos \phi \sin \omega t + I_2 \sin \phi \cos \omega t$$
$$= (I_1 + I_2 \cos \phi) \sin \omega t + I_2 \sin \phi \cos \omega t$$
$$= (I_1^2 + I_2^2 + 2I_1 I_2 \cos \phi)^{\frac{1}{2}} \sin \{ \omega t + \lrcorner I_2 \sin \phi / (I_1 + I_2 \cos \phi) \}$$

which is of the form $\quad i = A \sin (\omega t + \alpha),$

where A is the amplitude, α the phase angle, of the new sine curve of the same frequency as the constituents.

The particular case of quadrature is of interest and leads to the equation :—

$$i = (I_1^2 + I_2^2)^{\frac{1}{2}} \sin (\omega t + \lrcorner I_2/I_1).$$

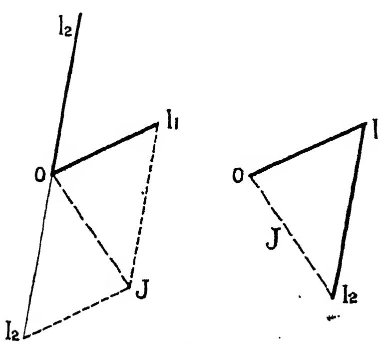


FIG. 67.

Difference of Two Equi-frequent Sines.

6. The taking of the difference by the vector construction is shown in Fig. 67. The resultant is seen to lag by more than a quarter period behind OI_2 and by rather less than a quarter period behind OI_1. Algebraically we write

$$i_1 - i_2 = (I_1^2 + I_2^2 - 2I_1 I_2 \cos \phi)^{\frac{1}{2}} \sin \{ \omega t - \lrcorner I_2 \sin \phi / (I_1 - I_2 \cos \phi) \}$$ which is of the form $\quad i = A' \sin (\omega t - \theta').$

Resolving Simple Harmonic Magnitudes into Components.

7. From the proposition that the sum of two sine curves of

equal period is another sine curve of the same period, it follows that one sine curve may be regarded as the sum of two or, in fact, any number of other sine curves of equal period. The process of finding either or both members of a pair that make up a given curve is called "resolving" and the members of the pair are called "components." The choice of pairs is infinite, and usually two conditions must be known or stated to make the task definite. For instance, in Fig. 68 the cranks OE and OI represent simultaneous values of the applied voltage and the current produced in a circuit. Suppose it is required to resolve the current into two components, of which one shall be in phase with the voltage and the other in quadrature. This is done by reversing the process of addition of vectors. The broken lines indicate the components I_p and I_w.

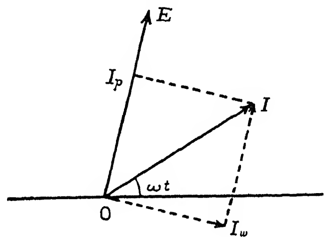


FIG. 68.

8. In most problems the two components sought are to be in quadrature with each other, but it should be remembered that the choice of even such pairs is infinite. It is only necessary to draw the circle on a given vector AB as diameter, and then to take any point P on the circle and join it to the ends of the vector. The two new vectors AP and PB together make up AB and are therefore one possible pair of its quadrature components.

Applied Emf. Potential Drop. Work. Signs.

9. The cranks and the laws governing them are used by electricians for calculation of, mainly, currents and voltage. They were used in physical optics, notably by Fresnel, more than a hundred years ago.

To remove possible ambiguities, especially regarding the signs of the voltages and the current in a circuit, it is necessary to lay down clearly the terminology to be used in this book. There is in Fig. 69 a graphic representation of the values of the electric potential existing at a particular instant along a circuit comprising a source of emf and a resistance. Suppose that on the left hand the potential relative to the earth is v_1. On passing

through the source of emf—which may be a machine or the secondary of a transformer, or, in the case of D.C., a voltaic or secondary battery—the potential relative to the earth rises to the value v_2. Passing along a conductor of negligible resistance we find it remains v_2, but along the resistance R it drops to the value v_3. The rise of potential $v_2—v_1$ is equal to e, the voltage generated by the machine. It is analogous to the rise of pressure imposed by a pump upon the water in the pipes of a water supply system. "Rise of potential" is an abbreviation for the phrase "increase of potential energy of a unit quantity of electricity" passing through the machine. Thus when i coulombs pass

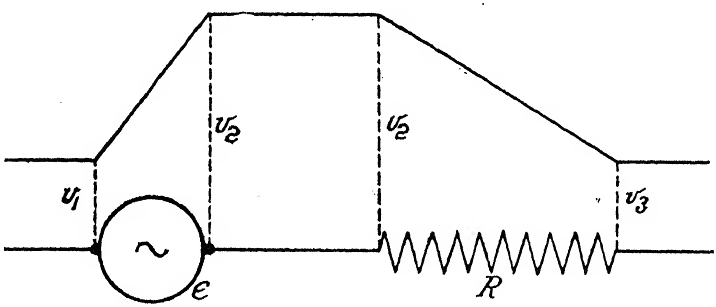


FIG. 69.

through the machine per second, that is to say, when the current is i amperes, potential energy is given to the electricity by the machine at the rate ei watts—energy which came ultimately from the prime mover driving the machine. It may be remarked that electromotive "force" is a misnomer; potential energy per unit charge is not force at all.

The potential drop along R is $v_2—v_3$, and this by Ohm's Law equals Ri. ("Potential drop" is merely an abbreviation for "the loss of potential energy per unit of electricity.") As the rate of passing of electricity through R is i coulombs per second, that is i amperes, the rate of loss of potential energy per second from the electric system is $(v_2—v_3)i$, which equals Ri^2. This is called Joule's Law. Since the potential drop along a resistance is positive when the current is positive and negative when the current is negative, being always in the same direction as the current, the rate of loss of energy in the resistance is always

positive. In the same way, the potential drop along an inductive coil or across a condenser will be written positive at any instant when it is in the same sense as the direction marked positive for the current, and *vice versâ*.

10. If the circuit of Fig. 69 be completed by a wire of negligible resistance, we must have $v_3 = v_2$, and therefore $e = Ri$. In words, the voltage applied by the generator is equal to or makes up for the potential drop along the resistance R; or the energy supplied to the circuit in the form of electric potential energy is equal in amount in a given time to that consumed in the resistance and disappearing as electric energy from the circuit to appear in the resistance as heat. It may seem at first sight confusing to equate a rise of potential to a fall of potential and to call both positive, though essentially opposite—but no difficulty arises if the sources of the voltage are distinguished carefully from the sinks. The drop of potential in the resistance and the rise of potential across the machine may be measured by a voltmeter and are called the terminal voltages of the machine and resistance.

Differential Coefficient of a Sine Curve.

11. It is clear from an inspection of a well-drawn sine curve that the steepest places are at the zero points and that the places where the curve is parallel to the datum line are the maxima and minima. If the gradient be determined at every point and the results plotted another sine curve is obtained. This tedious process can be replaced by aid of the crank diagram by applying the definition of differential coefficient. In Fig. 70 two positions of the rotating vector are shown separated by a very small time interval. Let dt be this interval, then the angle between the two positions is ωdt radian. Therefore the line joining the ends of the vectors is of length $I\omega dt$. Thus the rate of increase of the arc described from the datum position is $I\omega$ per second. The projection of this on the perpendicular to the datum line is $I\omega \cos \omega t$, since the join is perpendicular to either vector when

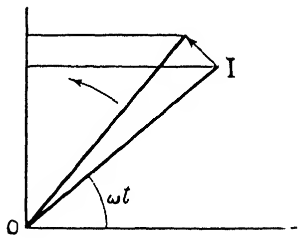


FIG. 70.

the angle and the time interval are made infinitesimal; that is, the differential coefficient of I sin ωt is $I\omega$ cos ωt. This written in the form of an equation becomes

$$DI \sin \omega t = I\omega \cos \omega t$$

where D means "rate of increase of" and stands for d/dt.

12. It is clear that i and its differential coefficient are repre-

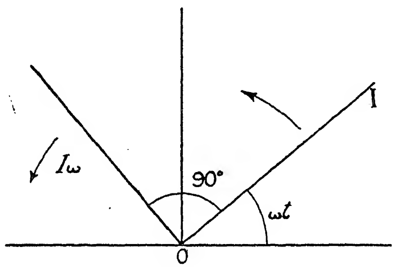


FIG. 71.

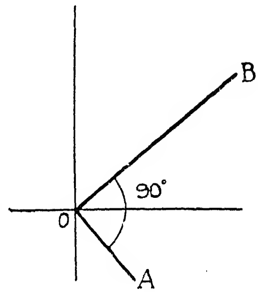


FIG. 72.

sented vectorially as in Fig. 71. The differential coefficient vector at the instant seized in the diagram has passed its maximum value but the current has not reached its maximum.

For the purposes of the crank diagram it is noteworthy that "differentiation" means "multiply the crank length by ω and rotate it forward 90°." This is all conveyed by the symbol D. In Fig. 72 OB represents to scale the differentiation of OA supposing $\omega = 2{\cdot}3$.

Reactance of a Coil.

13. The instantaneous value of the magnetic flux linkages produced when a current of magnitude $i = I \sin \omega t$ is passing through a coil of self-inductance L is Li. The rate of increase of the flux linkages is $D(Li)$ or, when L is invariable, LDi. The rate of increase of the linkages is therefore $L\omega I \cos \omega t$. According to the law of electromagnetic induction the increase of the linkages causes an emf in the coil equal to their rate of change and opposing the increase of the current. It is therefore called the back emf or the counter emf, or sometimes merely the inductive potential drop in the coil. Representing it by v, we have

$$v = L\omega I \cos \omega t = L\omega I \sin (\omega t + \tfrac{1}{2}\pi).$$

It leads the current by a quarter period and is of amplitude $L\omega I$. Thus we may write $V = L\omega I$ when thinking only of the magnitude and not the phase. The quantity $L\omega$ is called the reactance of the coil. In the equation $V = L\omega I$ it occupies a place analogous to R in the Ohm's law equation $V = RI$ and like resistance it is usually measured in ohms.

This result is placed on the crank diagram of Fig. 73, where the vectors are drawn so that the projections are i and v as defined by the above equations. This diagram may be obtained more directly by applying the geometrical rule for differentiating given at the end of the last paragraph.

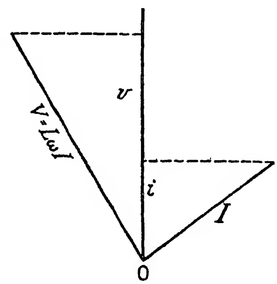


FIG. 73.

14. Reactance is so important in our subject that the above theory may with advantage be examined from another point of view. In Fig. 74 let us start with only the curve marked i drawn and suppose the current to be traversing a coil wound as a right-handed helix. Take the instant indicated by the point P on the curve; then the magnetic flux at the moment considered will thread the coil from left to right and will be increasing. By the law of induction the emf being induced is instantaneously in the direction from right to left. Therefore the potential at

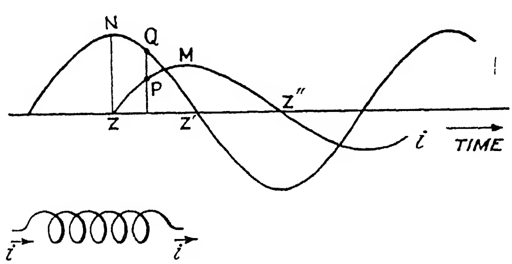


FIG. 74.

the left-hand end of the coil is greater than that at the right-hand end; or, in other words, the potential drop is from left to right, which is the direction of the current, and is therefore positive. Suppose this emf computed and set off to scale at Q. Now the steepest part of the current curve is at Z and therefore

the maximum induced potential drop will occur at the corresponding instant. Set up at Z the voltage ordinate ZN. Again, at M the current is neither increasing nor diminishing, the flux is stationary in value, and therefore the voltage being induced in the coil and the potential drop are zero at this instant. Mark this at Z′. The points NQZ′ determine a sine curve of the same period as the current. The amplitude has already been calculated to be $L\omega I$. It can be seen from the sketching in of the voltage curve that the drop along the coil is one quarter period ahead of the current through it. Calling ZM the first quadrant, MZ″ the second quadrant, and so on, it is to be noted that the potential drop is in the same direction as the current through the first quadrant, in the opposite direction through the second, in the same direction in the third, and in the opposite direction in the fourth. This is also clearly seen by rotating the cranks of Fig. 73, where the projections v, i are both on the same side of the datum line in two quarters of a period, and on opposite sides in the remaining quarters. If the circuit comprises only an inductance and a source of oscillatory emf, this applied emf must at every instant be equal to, and make up for, the potential drop in the inductance. Therefore when the potential drop is in the same direction as the current work is done in creating the magnetic field of the coil, and when it is in the opposite direction energy is recovered from the diminishing magnetic field. On the whole, as will be fully explained later, no work is done by the source of emf during the complete cycle.

Analogy to Magnetic Reactance.

15. A mechanical analogy may help at this point. The inductance coil may be likened to a flywheel on a shaft. And when a simple harmonic torque is applied to the shaft so as to make the flywheel rotate first in one direction and then in the other, the agent applying the torque feels a reaction due to the inertia of the flywheel. Provided there is little frictional resistance this inertia reaction is practically the whole of the opposition perceived, and its measure as a reactive force will be almost equal to the applied torque. If the motion is of large amplitude it is clear that the torque must be great; in fact the torque necessary to give any particular amplitude can easily be computed if the inertia of the wheel is known. The inertia of the flywheel

expressed as a "moment of inertia" corresponds to the inductance of the coil, the momentum of the flywheel to the reactance of the coil, and the rotational speed of the wheel to the electric current.

Numerical Example.

16. The reactance of a coil at any frequency is easily computed when its inductance and the frequency are given. For if f be the frequency and λ the wavelength in metres corresponding to the angular velocity ω, we have

$$L\omega = L \,.\, 2\pi f = L \,.\, 2\pi s/\lambda$$

where $s = 3 \times 10^8$, the velocity of electric waves in metres per second. Thus a coil with $L = 1$ millihenry has, at a wavelength of 1,885 metres, the reactance

$$\begin{aligned} L\omega &= 10^{-3} \,.\, 2\pi \,.\, 3 \times 10^8 \div 1{,}885 \\ &= 1{,}000 \text{ ohms.} \end{aligned}$$

Therefore to force an oscillatory current of 10 amperes through this coil at wavelength 1,885 metres will require a voltage of 10 kilovolts.

Area and Average Ordinate of Sine Curve.

17. The area of a portion of a given curve swept over by an advancing ordinate can be represented by another curve drawn on the same base and called the area curve. Any ordinate of the area curve represents the area of the given curve measured from some starting position of the moving ordinate

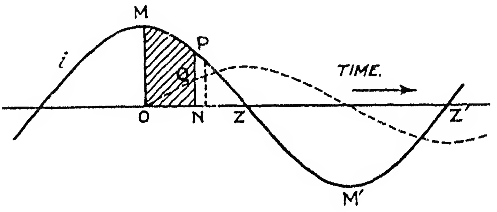


FIG. 75.

up to the common abscissa. Fig. 75 shows a cosine curve MPZM′Z′ and its area curve OQ, which is shown dotted. The ordinate QN of the area curve represents to scale the shaded area OMPN of the given cosine curve. If due allowance be made for the sign of the area—that portion appearing below the datum line being considered negative—it is seen easily that the area swept over as the ordinate NP advances will grow until Z is reached, will then begin to diminish, will become zero when M′ is reached (because two equal and opposite quadrants have been covered)

and will then become negative; thereafter its negative value increases until the position marked Z' is reached, and will then begin to diminish. All these variations, which can be deduced directly from inspection of the given cosine curve, are indicated by the ordinates of the dotted area curve.

Another connection between these two curves can be seen by imagining the ordinate PN to move forward to the position indicated by the dotted ordinate in Fig. 75 in the small time dt. In this step the area swept over is PNdt and this must be represented on the dotted curve by an increase in its ordinate. But the rise in the ordinate QN is by definition equal to the rate of increase of QN multiplied by dt, and therefore we deduce that the rate of increase of QN is equal to PN. In other words, the differentiation of the area curve yields the ordinates of the original curve, that is, the area curve may be obtained by "undifferentiating" the given curve. From the paragraph on the differentiation of sine curves we may also draw the conclusion that the undifferentiation, or integration as it is called, of a cosine curve with respect to time is another sine curve. Thus in symbols:—

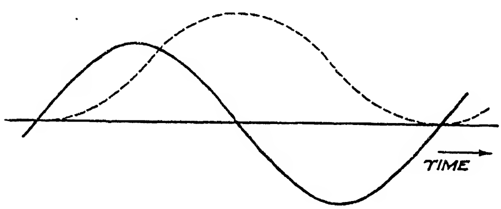


FIG. 76.

$$\int \cos \omega t dt = \frac{1}{\omega} \sin \omega t.$$

Since the result of differentiating a constant is zero, any constant may be added to the right-hand side of this equation, so that the complete formula might be written

$$\int \cos \omega t dt = \frac{1}{\omega} \sin \omega t + \text{constant}.$$

The meaning of this on a diagram may be illustrated, as in Fig. 76, by reckoning the area curve from the zero point instead of from the maximum point chosen in Fig. 75. At first sight, a very different area curve appears to have been obtained, but in fact this is the same curve as in Fig. 75 with a constant

added to every ordinate, or, as we may put it, lifted bodily a certain height above the datum line. The constant here is $1/\omega$. The above results include, of course, the integration of $\sin \omega t$, because the sine curve is merely the cosine curve pushed along the axis. The constant has to be determined afresh for every change of the starting point. In nearly all our electrical applications the constant is made zero by choice of the starting point.

Integration on the Crank Diagram.

18. For the purposes of the crank diagram, it is noteworthy that "integration" means "divide the cranklength by ω and rotate it backward 90°." Thus in Fig. 72 OA represents to scale the integration of OB, supposing ω to be about 2·3.

The operation of "integration" is represented symbolically by writing D^{-1} or $1/D$ in front of the symbol representing a vector. The notation is the logical consequence of writing D to represent "differentiation." It follows from § 12 that D^{-1} represents the operation of rotating a vector in the crank diagram backwards through 90° and dividing its length by ω.

Average Height of Sine Curves.

19. The average height of a complete alternation of a sine curve is obviously zero because for every ordinate above the datum line an equal similarly placed one can be found below the datum line. The average height of the positive ordinates, that is, of the ordinates in half an alternation, is obtained by dividing the base into the area, and is represented graphically by a line drawn at such a height above the base and parallel to it as to form a rectangle on the same base as the arch and of equal area. Thus we have for a curve of amplitude a,

$$\text{area of arch} = \int_0^{\frac{1}{2}T} a \sin \omega t dt$$

$$= 2a/\omega = aT/\pi$$

$$\text{and average ordinate} \quad = aT/\pi \div \tfrac{1}{2}T = (2/\pi)a.$$

Rectification.

20. Rectifiers of alternating current are devices possessing the property that when the emf acts in one direction a current flows through the device, and when it acts in the other direction a much smaller current—or, ideally, no current—is produced. The ideal rectifier would give a current with a graph such as is represented in Fig. 77, the arches being of sine shape. The average height of a whole number of these arches when taken over the

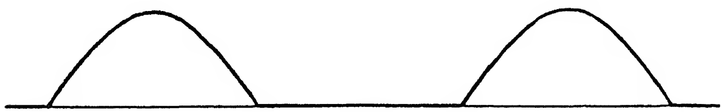


FIG. 77.

whole time is one-half of the average height of a single arch, that is to say, $1/\pi$ times the maximum height.

Some rectifiers of mechanical type can commutate, not merely rectify, and the graph of their current is as in Fig. 78. The average current in this case is $2/\pi$ times the amplitude. It is sometimes stated that the latter is more efficient than the former. This is not really the case—energy is not lost by suppression of the negative portion of a cycle by the non-commutating rectifier—

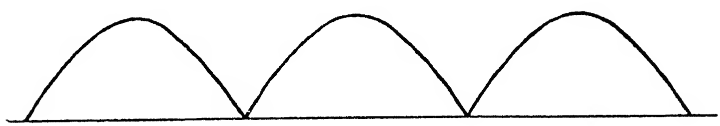


FIG. 78.

for the reason that as no current flows no work is done; but, though energy is not lost, the opportunity of adding to the unidirectional current is.

Reactance of a Condenser.

21. When a current $i = I \cos \omega t$ flows through a condenser, as represented in Fig. 79, the positive charge on the left-hand plate has a value q at a certain instant and the charge on the right-hand plate has the value $-q$. The current is represented as flowing into the one and out of the other to increase the value of q on both plates. The instantaneous voltage between the plates is q/C, if C is the capacitance, and the left-hand plate is at the higher potential. Thus there is at the instant defined by the

point P in the Figure a potential drop in the positive sense of the current, that is, a positive potential drop. The amount q of electricity on the left-hand plate, supposing there was none on it at the time $t = 0$, is obtained by adding up the current that has flowed into it; and this is

$$\int_0^t I \cos \omega t dt.$$

Thus by § 17, II. we have

$$q = (I/\omega) \sin \omega t,$$

and the amplitude of the alternating charge is $Q = I/\omega$. The voltage due to the charging of the condenser tends to oppose at

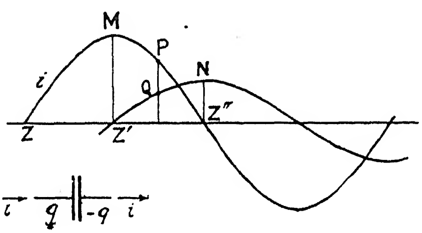


FIG. 79.

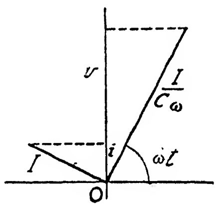


FIG. 80.

the instant chosen the further inflow of current to the condenser, and thus reacts against the applied voltage. The potential drop across the condenser is of amplitude

$$Q/C = I/C\omega = V, \text{ say,}$$

and thus the quantity $1/C\omega$ takes the place of resistance in the Ohm's Law equation, supposing the question of phase is not raised. This quantity is called the reactance of the condenser. The crank diagram is as in Fig. 80.

22. These results can be qualitatively checked by reasoning on the sine curve of Fig. 79. Suppose we start with only the current curve drawn in. At the instant determined by the point P the quantity of electricity on the left-hand plate, which has accumulated since the instant determined by the point M, may be represented by the ordinate under Q. This ordinate is really equal to the area of the current curve from the ordinate under M up to the ordinate under P. The area is, of course, zero at

Z', which is therefore a point on the curve of q. At the point Z'' the current changes sign, and therefore the charge q reaches at that moment its maximum value, which is marked off as the ordinate $Z''N$. Thereafter the q curve falls and is easily sketched in. It is evident from the curve that the charge, regarded as an alternating magnitude, lags a quarter period behind the current, and since the potential drop across the condenser is obtained by dividing the charge by a constant C, the potential drop across the condenser lags a quarter period behind the current. It will be seen also that this voltage is of the same sign as the current in the first quadrant $Z'Z''$, in the opposite direction in the second quadrant, in the same direction in the third and the opposite again in the fourth. These results are corroborated by reference to Fig. 80, where the projections of v and i are both on the same side of the datum line in two quarters of a period and on opposite sides in the remaining quarters.

23. If the circuit comprises only a condenser and a source of oscillatory emf, the applied emf must at every instant be equal to and make up for the potential drop in the condenser. Therefore when the potential drop is in the same direction as the current, work is done in creating the electric field of the condenser, and when the voltage is in the opposite direction energy is recovered from the diminishing electric field. On the whole no work is done by the source of emf.

Numerical Example.

24. The reactance of a condenser at any given wavelength is computed from the formula

$$1/C\omega = \lambda/2\pi sC.$$

Thus a condenser of capacitance $C = 1$ billifarad has, at a wavelength of 1,885 metres, the reactance

$$\begin{aligned} 1/C\omega &= 1{,}885 \div (2\pi \,.\, 3 \times 10^8 \,.\, 10^{-9}) \\ &= 1{,}000 \text{ ohms.} \end{aligned}$$

This result, like that of § 16, is worth remembering because it affords a basis for the rapid estimation of the reactance of any condenser at any wavelength. The frequency corresponding to 1,885 metres wavelength is 159,000 $\sim$ approximately.

Square of Sines. Root Mean Square.

25. In Fig. 81 is shown a sine curve of unit amplitude and also a dotted curve giving the squares of its ordinates. If the sine

curve is written $\sin \omega t$, the curve of the squares of the ordinates is to be written $\sin^2 \omega t$. But

$$\sin^2 \omega t = \tfrac{1}{2}(1 - \cos 2\omega t).$$

The formula shows that the dotted curve is also a sine curve, which has been lifted bodily above the datum line. Its own axis is sketched in. Now imagine the parts above the new axis cut off, turned over and fitted into the hollows. They will fit exactly because of symmetry, and we shall be left with an area between the limits indicated equal to the base multiplied by $\frac{1}{2}$. This area for a complete period is therefore $\frac{1}{2}T$ units. It is obvious that the average ordinate of the squared curve is of height $\frac{1}{2}$ a unit.

Suppose the original sine curve were $a \sin \omega t$; if this be squared we have $a^2 \sin^2 \omega t = \frac{1}{2}a^2(1 - \cos 2\omega t)$. To change the method,

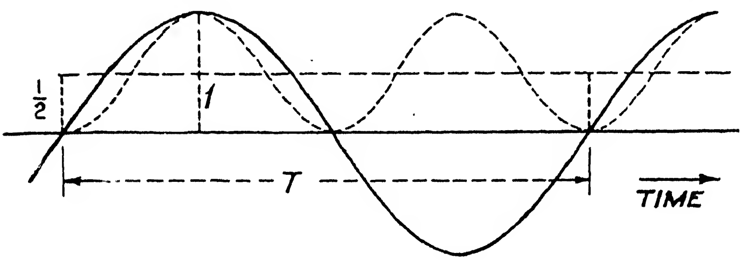


FIG. 81.

we may argue of this formula that the average of the 1 inside the bracket is 1, and the average of $\cos 2\omega t$ is zero because every cosine has as much negative as positive area inside a whole number of periods. Thus the average ordinate of the squared curve is $\frac{1}{2}a^2$.

26. The application of this to electrical cases is most easily seen by supposing an oscillatory current to pass through a resistance of R ohms. At any instant Ri^2 watts are being dissipated as heat, and therefore in order to know the energy dissipated in a complete period, we write

$$\int_0^T Ri^2 dt,$$

or, in words, we find the area of the curve of i^2 multiplied by R. From what appears above, this will be $\frac{1}{2}RI^2$. Suppose now the oscillatory current replaced by a steady current which

can be adjusted to dissipate the same amount of energy per second as the oscillatory current. Let its magnitude be I_0. Then

$$RI_0^2 = \tfrac{1}{2}RI^2.$$

Thus
$$I_0 = \sqrt{\tfrac{1}{2}}I = 0{\cdot}707I.$$

This equivalent steady current is called the effective, or the virtual, or the root mean square, value of the oscillatory current, always supposing this last to be of perfect sine form.

The symbol A will be used for the effective (or rms *) value of an oscillatory current of amplitude I whenever it is necessary to use the two symbols in one equation, but as a rule it is possible to take the symbols in electrical equations as relating either to amplitudes or to rms values throughout the mathematical work. The same statement holds good with regard to crank diagrams: they may be supposed to have any I marked on them expressed as an rms value or as an amplitude. It is merely a matter of scale in reading the diagram.

27. It is evident from the preceding paragraphs that a hot-wire ammeter is by nature well adapted for the task of measuring alternating currents, since the extension of its sensitive wire is a measure of the square, and not the first power, of the current. And analogously, the forces in an electrostatic voltmeter are proportional to the square of the potential difference applied to the terminals. Thus if an alternating voltage of amplitude V is applied to such an instrument and then an unvarying voltage of amount V_0 applied and adjusted to give the same reading, we shall have $V_0^2 = \tfrac{1}{2}V^2$, whence $V_0 = 0{\cdot}707V$. But it should be noticed that either instrument may be graduated in rms or in amplitudes.

The Product of Sines.

28. The product of two equifrequent sine curves is another sine curve. This can be shown by taking two sine curves as in Fig. 82 and carrying out the multiplication of ordinates to give the new curve drawn dotted in that Figure. Suppose one of the sine curves to be $I \sin \omega t$ and the other to be $E \sin (\omega t + \phi)$. Then from elementary trigonometry we have the product

$$\begin{aligned} ei &= EI \sin(\omega t + \phi) \sin \omega t \\ &= \tfrac{1}{2}EI \left\{ \cos\phi - \cos(2\omega t + \phi) \right\}. \end{aligned}$$

* We shall use the abbreviation rms for "root mean square."

This proves the above statement. It shows also that the product curve is of double the frequency of the original curves.

29. To obtain the area of this product curve we first notice that there must be an axis of symmetry. This is drawn as a broken line in the Figure. The trigonometry shows that the product curve is lifted a distance of $\cos\phi$ above the axis of the original curves. Therefore the area on a length of base corresponding to a complete period is $\frac{1}{2}T\ \cos\phi \times EI$. This follows from the Figure by imagining the upper arches to be cut off and turned over to fill the hollows below the dotted axis, or it may be seen from the trigonometrical expression by noticing that the average of $\cos\phi$ is $\cos\phi$ and of the variable cosine portion of the

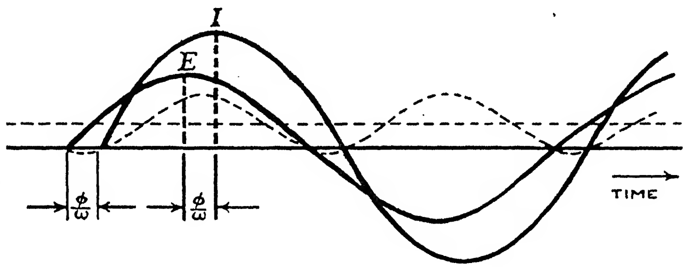


FIG. 82.

expression is zero. Thus the average ordinate of the product curve is $\frac{1}{2}EI\ \cos\phi$.

Power and Power Factor.

30. This result has an important electrical application. For we know that the instantaneous power being expended in any branch traversed by a current $i = I\ \sin\omega t$ when the terminal voltage is $e = E\ \sin(\omega t + \phi)$ is ei and is as expressed in § 28.

The total work done in one period is equal, therefore, to the area of the product curve, namely, $\frac{1}{2}EIT\ \cos\phi$. Therefore the energy expended in a second must be

$$P = \tfrac{1}{2}EI\ \cos\phi$$

where E and I are both amplitudes. But as the amplitude is E the rms value of the voltage is $\sqrt{\frac{1}{2}}E$, and as I is the amplitude of the current its rms value is $\sqrt{\frac{1}{2}}I$. Therefore the energy expended in a second is

$$\text{rms terminal voltage} \times \text{rms current} \times \cos\phi$$

for sine curves.

The multiplier $\cos\phi$ is called the power factor.

In the case of pure resistances the terminal voltage will be in phase with the current, and therefore ϕ will be zero, and $\cos\phi$ will be unity. The result is that the rate of expenditure of power is in this case

rms terminal voltage $\times$ rms current.

31. These results can be put on the crank diagram. In Fig. 83 let us suppose the E and I vectors to be on the amplitude scale. Then of the two components the one marked I_p, which is taken in phase with E, is equal to $I\cos\phi$. Therefore, from the equation above, we obtain

$$P = \tfrac{1}{2}EI_p$$

for the average electrical power being developed in the part of the circuit to which the diagram refers. This inphase component of the current is sometimes called the "load" component, sometimes the "power" component. The other component, which in the long run does no work with E, is called the "idle" or "wattless" component of the current.

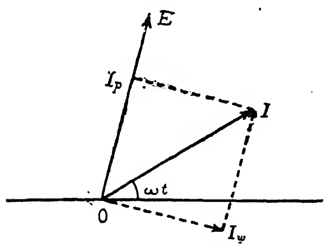


FIG. 83.

It will be noticed that the amplitude of the load component of the current is equal to the value of the current at the moment the voltage is a maximum, and that the amplitude of the wattless component of the current is equal to the value of the current at the moment the voltage is zero. This statement remains true if the words current and voltage are interchanged.

OVERTONES OR HARMONICS

32. When by any means a sine emf and a sine current of different frequency are produced in the same circuit the former performs no work on the latter if the reckoning be averaged over a large number of periods of either. Circumstances like this arise when a current of one frequency is passed through a machine or apparatus generating another frequency. The proposition is easily seen to be true when one frequency is a multiple of the other

—for simplicity, take the case when the current is treble the frequency of the emf so that

$$e = E \sin \omega t,\ i = I \sin (3\omega t + \theta) \tag{1}$$

The instantaneous rate of working is ei and the work done in one period of the slower alternation is

$$\int_0^{2\pi/\omega} ei dt = EI \int_0^{2\pi/\omega} \sin \omega t \sin (3\omega t + \theta) dt$$

$$= \tfrac{1}{2}EI \int_0^{2\pi/\omega} [\cos (2\omega t - \theta) - \cos (4\omega t + \theta)]dt \tag{2}$$

$$= \tfrac{1}{2}EI \left[\frac{\sin (2\omega t - \theta)}{2\omega} - \frac{\sin (4\omega t + \theta)}{4\omega}\right]_0^{2\pi/\omega}$$

$$= 0.$$

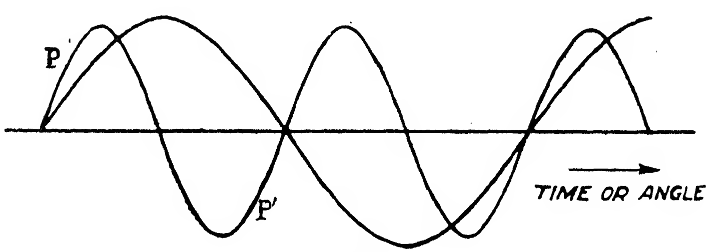


FIG. 84.

For the case when one frequency is an even multiple of the other algebra is not needed; for, taking the case of Fig. 84, it is clear that the product of the ordinates between 0 and $\frac{1}{2}\pi$ is positive, while from $\frac{1}{2}\pi$ to π it is negative. Moreover, for every product of positive ordinates such as those at P an equal pair giving a negative product can be found at P′. Therefore the total work done in a half period of the fundamental is nil. In general, since the product of any two sines is expressible as the difference of two cosines, and since also the area of each of these cosines is zero for its own complete period, we conclude that the total area of the cosine curves will certainly be zero when a time is taken that is an integral multiple of both their periods. In practice all the instances of greatest importance are those in which one of

the given frequencies is an integral multiple of the other—one is a fundamental and the other a harmonic—and then the area of the product of the two given curves is zero within a single period of the fundamental. For we have by trigonometry, taking the harmonic as of frequency m times the fundamental :—

$$2 \sin \omega t \sin (m\omega t + \theta)$$
$$= \cos \{(m-1)\omega t + \theta\} - \cos \{(m+1)\omega t + \theta\} \qquad (3)$$

In one period of the fundamental the angle ωt ranges from 0 to 2π, which causes one of the above cosines to go through $m - 1$ complete periods and the other through $m + 1$ complete periods. Each cosine therefore yields no area. Hence a fundamental emf cannot do any electrical work with a harmonic current and conversely. But work is, of course, done by an emf of the harmonic frequency acting upon the current of the same frequency ; its amount being, as proved already, $\frac{1}{2}EI \cos \phi$ in our customary notation.

Fourier Analysis.

33. These facts enable us to determine in any ordinary periodic function of time the amplitude of the fundamental and each of its harmonics. This mathematical process is called "Fourier Analysis" after the mathematician who showed that most physical periodic functions of time may be regarded as composed of fundamental and harmonics.

Let it be required to find the amount of the harmonic of five-fold frequency in the current curve of Fig. 85. Assume it to be $I_5 \sin (5\omega t + \theta_5)$ where both I_5 and θ_5 are unknown. It is best to write this in the equivalent form

$$I_5 \cos \theta_5 \sin 5\omega t + I_5 \sin \theta_5 \cos 5\omega t. \qquad (4)$$

Imagine a voltage $E \sin 5\omega t$ imposed on the circuit. This sine curve is drawn in the Figure and the drawing board task consists in multiplying ordinates and finding graphically the area of the product curve thus arrived at. Call this area S_5. Since no work is done between the applied emf and the other partials, the area so determined gives information about the quintuple harmonic only. Actually the work done per second must be equal to $\frac{1}{2}EI_5 \cos \theta_5$ with the sine component assumed above, and no work is done with the cosine component because this is in

quadrature with the applied emf. The work done in a period of the given curve is

$$\tfrac{1}{2}TEI_5 \cos\theta_5$$

where $T = 2\pi/\omega$. Equating this calculated area to that determined graphically, we obtain

$$I_5 \cos\theta_5 = 2S_5/TE. \tag{5}$$

All the terms on the right of this equation are known.

Now suppose the applied emf to be $E \cos 5\omega t$. Draw this on the board, take the products of the ordinates and find the area

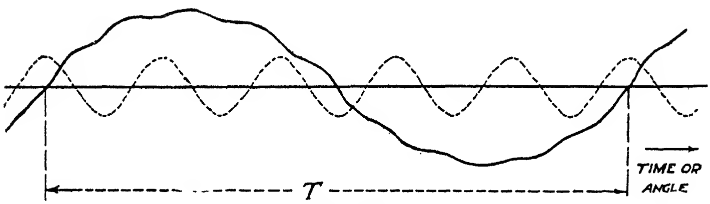


FIG. 85.

over a complete period of the given curve. Let this area be C_5. Then by reasoning similar to that above we get

$$I_5 \sin\theta_5 = 2C_5/TE \tag{6}$$

From these two equations the two unknowns can be determined thus :—Squaring the equations and adding them leads to

$$I_5 = 2(S_5^2 + C_5^2)^{\frac{1}{2}}/TE. \tag{7}$$

Dividing gives

$$\tan\theta_5 = C_5/S_5. \tag{8}$$

In this formula E might as well be taken as 1 volt.

Similar formulæ hold for harmonics of any order. It should be mentioned that sometimes there is a D.C. component in a given periodic current, and that this is determined by measuring the area of a complete period of the given curve and dividing by the complete period—which means merely finding the average ordinate.

34. Occasionally the given curve is of simple form and expressible by equations ; then the finding of the harmonics may be done by the integral calculus instead of graphically. Take, for

instance, the curve of a regularly interrupted current as given in Fig. 86. It has two equations, namely:

$$i = I \text{ from } t = 0 \text{ to } t = \tfrac{1}{2}T$$
$$i = 0 \text{ from } t = \tfrac{1}{2}T \text{ to } t = T.$$

The first thing to do is to find its direct current component. Clearly the area of a whole period is $I \cdot \frac{1}{2}T$, and the average ordinate is $\frac{1}{2}I$. The dotted line in the figure thus represents the D.C. component.

The fundamental will be got by multiplying first by $\sin \omega t$ (the voltage being taken as unity) and integrating over a period. From $t = 0$ to $\frac{1}{2}T$ we have to integrate $I \sin \omega t$, and from $t = \frac{1}{2}T$ to T we integrate $0 \cdot \sin \omega t$. We get $2I/\omega$ for the former and 0 for the latter, and therefore $S_1 = IT/\pi$. Now we apply the emf $\cos \omega t$, multiply it into the given curve and integrate. In-

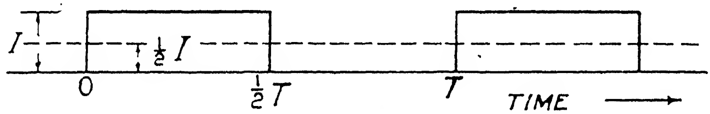


FIG. 86.

spection of the sketch shows that this area is zero, therefore $C_1 = 0$. Equations (7) and (8) now give

$$I_1 = 2I/\pi$$
$$\tan \theta_1 = 0.$$

The harmonics can be got one by one, but it is better to get a general formula. Apply the emf $\sin m\omega t$ to the current curve, multiply and integrate. We get

$$S_m = \int_0^{\frac{1}{2}T} I \sin m\omega t dt + \int_{\frac{1}{2}T}^{T} 0 \cdot dt$$

$$= I\left[\frac{-\cos m\omega t}{m\omega}\right]_0^{\frac{1}{2}T} = \frac{I}{m\omega}(1 - \cos m\pi).$$

The cosine component gives $C_m = 0$. Thus

$$I_m = 2S_m/T$$
$$= \frac{I}{m\pi}(1 - \cos m\pi).$$

Taking the harmonics in turn this formula gives

$$I_2 = 0,\ I_3 = 2I/3\pi,$$
$$I_4 = 0,\ I_5 = 2I/5\pi, \text{ etc.}$$

Thus the even numbered harmonics are absent and only the odd

numbered ones, together with a D.C. component, exist. The current may now be written

$$\tfrac{1}{2}I + \frac{2}{\pi}I\left[\sin \omega t + \tfrac{1}{3}\sin 3\omega t + \tfrac{1}{5}\sin 5\omega t + \ldots\right]$$

A series such as this is called a Fourier Series.

Guiding Rules.

35. In the harmonic analysis of periodic curves labour may often be lessened by inspection of the curve after its mean line has been drawn. First notice whether the negative arch is or is not a repetition of the positive arch (with sign changed). All sine and cosine curves of frequency an odd multiple of the fundamental have the property of repeating at the half period with change of sign, and therefore any departure from this condition in a given curve implies the presence of even-fold harmonics. Next examine whether there is symmetry on passing towards the right and left hand from each of the zeroes of the curve; in this case fundamental and harmonics can be written as sines passing through all the zeroes. Then look for right and left-hand symmetry at the mid-ordinate of an arch; the cosines are suggested by this property. All these statements can be deduced from a few sketches of a fundamental sine and its harmonics. They are tabulated below. An entry such as "cosines only" means that if the fundamental be written as a cosine by taking the origin suitably (the middle of an arch), all the harmonics prove to be cosines also.

	Negative arch a repetition of positive, therefore even-fold harmonics absent.	Negative arch not a repetition of positive, therefore even-fold harmonics present.
Right and left symmetry about mid-period, with change of sign.	Sines only.	Sines only.
Right and left symmetry about middle of each arch.	Can be written as sines only, or cosines only.	Cosines only.

Besides these rules the labour may often be shortened by knowledge of the circumstances in which the voltage or current is generated. For instance, in ordinary alternators the positive and negative waves are generated under identical conditions and therefore there are no even-fold harmonics. In star connected three-phase machines there is no triple harmonic nor any multiple of the triple frequency.

OPERATORS AND CRANK DIAGRAMS OF VARIOUS CIRCUITS

Resistance Operators.

36. It has already been explained that a simple harmonic oscillatory current is represented algebraically by

$$i = I \sin \omega t,$$

that its rate of increase Di is

$$Di = I\omega \cos \omega t,$$

and that integration or accumulation is the same as undifferentiation and is therefore

$$\frac{1}{D} i = -\frac{I}{\omega} \cos \omega t$$

plus a constant, that is to say, a non-oscillating term that can usually be ignored in problems about oscillations.

By applying this algebra to certain electrical laws we saw that we could express the difference of potential at the terminals of a resistance, an inductive coil, or a condenser, when a current i is passed through it, by sines and cosines. These terminal potential differences may also be written very briefly in the form

$v_R = Ri$	for a resistance ;
$v_L = LDi$	for a coil ;
$v_C = (1/CD)i$	for a condenser.

The symbols of mathematical operation LD and $1/CD$ occupy an analogous position to R in the first of the three equations and have been called by Heaviside "resistance operators." The differentiator D itself (and consequently the resistance operators) can be shown to obey some of the rules of algebra just as if D were a quantity capable of having a numerical value. For

instance, in its application to sines and cosines, which is the only application we need, we have

$$D \sin \omega t = \omega \cos \omega t$$
$$D(D \sin \omega t) = - \omega^2 \sin \omega t.$$

It is natural to write this last in the form

$$D^2 \sin \omega t = - \omega^2 \sin \omega t$$

where the index over D does not mean repeated multiplication as it would for a numerical quantity, but is agreed upon as meaning repeated differentiation. Thus

$$D^3 \sin \omega t = D(- \omega^2 \sin \omega t)$$
$$= - \omega^3 \cos \omega t$$

and
$$D^4 \sin \omega t = \omega^4 \sin \omega t,$$

and so on.

Again $1/D$ can be written D^{-1} and we have seen that

$$D^{-1} \sin \omega t = - \omega^{-1} \cos \omega t.$$

Therefore
$$D^{-2} \sin \omega t = - \omega^{-2} \sin \omega t,$$

and also

$$D \,.\, D^{-1} \sin \omega t = D(- \omega^{-1} \cos \omega t)$$
$$= \sin \omega t.$$

But by the rules of indices $D \,.\, D^{-1} = D^0$ and therefore we see

$$D^0 \sin \omega t = \sin \omega t$$

or D^0 means " no differentiation."

37. From the above differentiations and integrations an important rule emerges: the effect of D^2 on $\sin \omega t$ (or on $\cos \omega t$) is obtained by replacing it by $- \omega^2$ wherever it occurs. For example in $D^{-2} \sin \omega t$ we are dealing with $1/D^2$, and on replacing D^2 by $- \omega^2$ we get $- 1/\omega^2$, the same as obtained by actual differentiation. But the rules may be applied in more complicated cases. For example

$$D(aD + bD^{-1}) \sin \omega t = (aD^2 + bD^0) \sin \omega t$$
$$= (- a\omega^2 + b) \sin \omega t$$

as may be proved by direct differentiation. Again such an operator as $(D + a)^{-1}$, which is a sort of integration, can be simplified by a process analogous to " rationalisation of denominator " in algebra, in which the endeavour is to reduce operators in the denominator to squares so that they can be replaced by $- \omega^2$. Thus

$$
\begin{aligned}
(D+a)^{-1} \sin \omega t &= \frac{1}{D+a} \sin \omega t \\
&= \frac{1}{D+a} \cdot \frac{D-a}{D-a} \sin \omega t \\
&= \frac{D-a}{D^2-a^2} \sin \omega t \\
&= -\frac{1}{\omega^2+a^2}(D-a) \sin \omega t
\end{aligned}
$$

which reduces the complicated integration to a single differentiation.

38. Again from the equations of § 36 we obtain

$$D^{-1} \sin \omega t = -\omega^{-1} \cos \omega t = -\omega^{-2} D \sin \omega t$$

indicating that $\qquad D^{-1} = -\omega^{-2} D$

that is $\qquad 1/D = -D/\omega^2.$

This, of course, is only another way of writing the equivalent $D^2 = -\omega^2$. As an example of its application, consider the effect of the resistance operator $LD + R + 1/CD$ on $\sin \omega t$. We have

$$
\begin{aligned}
LD + R + 1/CD &= LD + R - D/C\omega^2 \\
&= R + (L - 1/C\omega^2)D \\
&= R + (L\omega - 1/C\omega)D/\omega \\
&= R + XD/\omega
\end{aligned}
$$

where, as is often done for brevity, we have put

$$X \equiv L\omega - 1/C\omega.$$

The importance of this result will be seen later. It includes obviously, the formula

$$LD + 1/CD = XD/\omega$$

wherein a combined differentiation and integration are reduced to a single differentiation.

(The reader not interested in the fuller applications of resistance operators should read the definition of Impedance and pass at once to § 49.)

Impedance as an Operator

Definition of Impedance.

39. We have seen in the last section that inductance coils and condensers, like resistances, hinder or impede the passage of sine current each in its own way. By the use of the conception reactance we have expressed this hindrance in ohms,

though reactance differs profoundly from resistance in that it is a kind of hindrance that does not consume energy but introduces phase changes, as already explained. We shall often meet in succeeding pages with circuits that comprise both resistance and reactance, and therefore simultaneously offer both kinds of hindrance. The result is called the impedance of the circuit. In calculating it both magnitude and phase have to be considered. The resistance operator method makes this task easy.

40. The whole expression $R + XD/\omega$ represents the impedance of a circuit symbolically. It appears typically in equations like the following :—

$$e = (R + XD/\omega)I \sin \omega t$$

which is the full extension of Ohm's Law to alternating current. We have

$$\begin{aligned} e &= I(R \sin \omega t + X \cos \omega t) \\ &= (R^2 + X^2)^{\frac{1}{2}} I \sin (\omega t + \angle X/R) \end{aligned}$$

or more shortly

$$e = ZI \sin (\omega t + \phi).$$

In electrical language this states that the emf needed to send an alternating current of amplitude I through a circuit possessing resistance R and reactance X is of amplitude ZI and leads the current by the angle whose tangent is X/R. Taken to the crank diagram we should draw a vector to represent I, multiply it by Z, and then rotate it forward by the amount of the phase angle ϕ, in order to get the vector representing the necessary emf.

For the sake of brevity we shall use (Z, ϕ) to represent $R + XD/\omega$, that is to represent both the magnitude Z and the angle ϕ of the impedance. In fact (Z, ϕ) will be used to convey all the properties of the impedance, merely as a piece of shorthand.

Addition of Impedance Operators.

41. The meaning of such an expression as

$$(Z, \phi) + (Z', \phi')$$

is easily arrived at by writing the operators in full.

It is $\quad R + XD/\omega + R' + X'D/\omega$

or $\quad R + R' + (X + X')D/\omega.$

This is an impedance of magnitude

$$Z_2 = \{(R + R')^2 + (X + X')^2\}^{\frac{1}{2}}$$

with a phase angle whose tangent is

$$(X + X')/(R + R'),$$

say the angle ϕ_2.

It will be found that the geometry of a crank diagram will give precisely the same result; examples appear below.

The difference of two impedance vectors is treated in the same way.

Product of Impedance Operators.

42. The product of two impedances sometimes arises. Consider the product $(Z, \phi)\,(Z', \phi')$ and let us interpret it by supposing it applied to a vector. We take the meaning of the product to be: first operate by one and then operate on the result by the other. Now operation by (Z', ϕ') means multiply the given vector by Z' and rotate it forward through the angle ϕ'. To operate on this new vector by (Z, ϕ) means multiply by Z and rotate the crank forward by the angle ϕ. In symbols

$$\begin{aligned}(Z, \phi), (Z', \phi') \sin \omega t &= (Z, \phi)Z' \sin (\omega t + \phi') \\ &= ZZ' \sin (\omega t + \phi + \phi').\end{aligned}$$

The final result is that we have an impedance of magnitude ZZ' and with a phase angle $\phi + \phi'$ or, briefly

$$(Z, \phi)(Z', \phi') = (ZZ', \phi + \phi').$$

Thus in a product of two impedance vectors, the magnitudes are multiplied but the phase angles added.

The quotient of two impedance vectors or operators may be discussed in the same way. Division implies negative rotation, therefore

$$\frac{(Z, \phi)}{(Z', \phi')} \sin \omega t = \frac{Z}{Z'} \sin (\omega t + \phi - \phi').$$

In words, the magnitudes divide but the phase angles subtract.

Reduction of more complicated Expressions.

43. Once D^2 has been replaced by $-\omega^2$ it has fulfilled its function as an operator and finished its career. In consequence any expression like

$$aD^8 + bD^7 + cD^6 + eD^3 + f$$

can be reduced to

$$a(-\omega^2)^4 + b(-\omega^2)^3D + c(-\omega^2)^3 + e(-\omega^2)D + f$$

or

$$a\omega^8 - c\omega^6 + f - (b\omega^6 + e\omega^2)D,$$

which is of the form

$$A + BD.$$

This reminds one of the general expression for an impedance, which is

$$R + (X/\omega)D,$$

and therefore we reach the important conclusion that a complicated circuit that gives an operator like the one we started from can be represented by a simple impedance containing a single resistance and a single reactance—the latter being an inductance or a capacitance according to the sign B happens to have.

44. Moreover, a fraction made up of two expressions like the given one may have both its numerator and denominator reduced to first degree binomial form. Then the denominator can be " rationalised " as explained in § 37; which will leave the completed fraction again in the form $A + BD$. Now all circuits comprised of concentrated inductance and capacitance, however complicated, have operators that reduce in the above manner. And therefore we have the very important theorem that every such circuit is capable of being expressed as and represented by a single resistance and a single reactance connected in series. The value of the equivalent resistance and of the equivalent reactance will depend on ω, that is, on the frequency.

Use of "Imaginary" and "Complex" Numbers.

45. It may be noticed that since $D^2 = -\omega^2$ we may write $D = \omega \sqrt{(-1)}$. We have already explained that on a crank diagram D means " multiply by ω the vector operated upon and rotate it anticlockwise through a right angle," and therefore we conclude that $\sqrt{(-1)}$ must be capable of interpretation as rotation through a right angle anticlockwise. This interpretation was brought into prominence by Argand more than a century ago and has since played an important part in the study of the " imaginary " quantity, in the mathematical physics of vibrations (especially light), and, much later, in alternating current theory. As it would necessitate a whole section to explain the rules of the algebra of complex numbers, as numbers with an imaginary part are called, we shall not adopt that method of analysing oscillatory circuits. We shall instead be content with the resistance operator method, which is indeed a perfect equivalent. The sole disadvantage of this procedure is that although we are able to substitute for D^2 by the rule $D^2 = -\omega^2$, we cannot get rid of plain D except by an actual differentiation. For

$$D \sin \omega t = \omega \cos \omega t,$$

and this cannot under any circumstances be written as a multiple of $\sin \omega t$. The cosine shows, however, the change in direction

through a right angle brought about by D when it acts on the vector $\sin \omega t$, and we must always keep in mind, on transferring operational equations to crank diagrams, that a surviving D in an expression carries in itself both multiplication by ω and rotation forward through a right angle.

It may be remarked, by the way, that D^2 carries in itself two successive quadrantal rotations; and because two such rotations are equivalent to a reversal in direction we are able to convey all the directional effect of D^2 by writing a minus sign in front of the numerical part of its equivalent.

Steady Oscillation. Transient Oscillation.

46. When a sine emf is suddenly applied to or removed from circuits of the kind alluded to in the preceding paragraphs the oscillatory current takes time to build up to a steady state or to die down from a steady state. Nearly all the practical problems of continuous wave telegraphy, however, are such that the preliminary or final variable states may be disregarded; and in consequence of this we look upon steady state solutions of our problems as fully sufficient for practical purposes. The solutions obtained by putting $-\omega^2$ for D^2 in resistance operator equations are steady state solutions—they ignore the building up or dying down of the oscillatory current just as these same phenomena are often ignored in alternating current engineering.

The discussion in § 38 of the resistance operator of a circuit containing inductance, resistance and capacitance in series is merely a preparation for solving the differential equation

$$(LD + R + 1/CD)i = E \sin \omega t$$

without regard to the building-up interval. The corresponding diagram is merely another way of solving this same equation. This equation if differentiated and divided by L yields

$$\left(D^2 + \frac{R}{L}D + \frac{1}{LC}\right)i = \frac{E\omega}{L} \cos \omega t$$

which is also solved by the methods mentioned.

47. In spark telegraphy the typical differential equation is, on the other hand,

$$(LD + R + 1/CD)i = 0,$$

for when the condenser is charged and the emf suddenly removed by the spark jumping the gap and completing the circuit, the applied emf is zero. In spark telegraphy the transient damped

oscillation is obviously the important thing. The methods discussed in this book will not solve this equation or the equivalent equation

$$\left(D^2 + \frac{R}{L}D + \frac{1}{LC}\right)i = 0.$$

This statement is easily verified by trying the substitution $D = \omega\sqrt{-1}$, when it will be seen that no possible value of ω can make the left-hand number of the equation zero. To find the proper substitution for D is easy, however. For on solving

$$D^2 + \frac{R}{L}D + \frac{1}{LC} = 0$$

as a quadratic equation in D we obtain

$$D = -\frac{R}{2L} \pm \left(\frac{1}{LC} - \frac{R^2}{4L^2}\right)^{\frac{1}{2}}\sqrt{-1}$$

or say
$$D = -b + p\sqrt{-1}.$$

It would take too much space to explain that this result indicates that the solution will be of form

$$i = I\epsilon^{-bt}\sin(pt + \alpha),$$

but the reader may verify the statement by trying the effect of differentiation upon this equation. The quantity b is called the decay coefficient of the circuit, and the quantity p the natural or free angular velocity. When the resistance of the circuit is negligible the decay coefficient vanishes and the natural angular velocity becomes

$$p_0 \text{ or } \omega_0 = \frac{1}{LC}$$

Clearly, in general

$$p^2 = \omega_0^2 - b^2.$$

The decaying oscillation here described may be represented graphically by a crank whose length diminishes exponentially, whose end therefore describes a logarithmic spiral in place of the circle of alternating current crank diagrams.

When a sine emf is suddenly applied to the circuit just considered the current builds up by means of a damped oscillation of which the principal part can be represented by the equation for i given above. When the sine emf is suddenly removed the current dies down according to this same equation.

IMPEDANCE OF VARIOUS CIRCUITS

48. Impedance has been defined in § 39, and from this definition the impedance of any circuit can be calculated by aid of the operators discussed in §§ 36 to 47. But in most of the paragraphs to follow, the impedance of typical circuits will be discussed by aid of crank diagrams because this way is perhaps more instructive than the method of resistance operators. But the reader interested in resistance operators is recommended to supplement the crank diagram method by the operator method for himself.

SERIES ARRANGEMENTS

Resistance and Inductance in Series.

49. Let R represent the resistance in ohms and L the inductance in henrys of the conductor and the coil in Fig. 87, and let a current I (rms or amplitude value) be traversing both. The potential drop along the conductor is RI, and that along the coil is $L\omega I$, speaking in rms or amplitude values as the case may be. To set these on the crank diagram we take any random direction for I, mark RI along it to scale as OA, and set off AB a right angle forward to represent $L\omega I$ on the same scale. Both OA and AB will then be in volts. Thus if $R = 800\ \Omega$, and $L = 1$ mH, while $\omega = 10^6$ and $I = 1$ A, we shall have OA $= 800$ V and AB $= 1{,}000$ V.

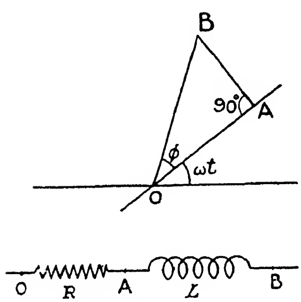


FIG. 87.

These are the readings of voltmeters connected across the terminals of the resistance and the coil. At the same time a voltmeter connected across both would give a reading corresponding to the length OB in the crank diagram. This length is given by

$$\mathrm{OB}^2 = \mathrm{OA}^2 + \mathrm{AB}^2$$

or

$$\mathrm{OB} = \sqrt{(800^2 + 1{,}000^2)}$$
$$= 1{,}280 \text{ volts.}$$

The P.D. across the two parts is thus less than the sum of the

separate P.D.'s on account of the phase difference that exists between them.

It need hardly be said that as the cranks rotate their perpendicular projections give sine curves, and the pair of curves representing the component voltages add together to give the third or resultant voltage. Of the resultant voltage OA is the power component, AB the wattless component.

50. Suppose the circuit to comprise only the resistance, the inductance and the source of emf of amplitude E, then OB will represent E to scale, and the right-angled triangle immediately gives

$$E = \sqrt{(R^2 + L^2\omega^2)} \,.\, I.$$

The diagram points out also that the voltage leads the current by the angle marked ϕ, and that

$$\tan \phi = L\omega/R.$$

The former equation tells us that the current is related to the voltage by a rule analogous to Ohm's Law for unvarying currents, but the resistance in the Ohm's Law equation is here replaced by $(R^2 + L^2\omega^2)^{\frac{1}{2}}$. This quantity has already been called the impedance of the circuit and given the symbol Z. Thus

$$Z = \sqrt{(R^2 + L^2\omega^2)}.$$

These two equations, or their equivalents, are always both needed to describe fully the combined effect of resistance and reactance in a circuit.

Application to Measurement of Constants of Coil.

51. The impedance of a coil at any frequency $\omega/2\pi$ may sometimes be determined with sufficient accuracy in the following simple manner. Let L and R be the unknown constants of the coil. Connect in series with the coil a known non-inductive resistance R' and suppose a current of I amperes and of frequency $\omega/2\pi$ to pass. If a delicate voltmeter of small capacitance be available, measure the terminal voltages indicated at FG and GH in Fig. 88. Let these be V^1

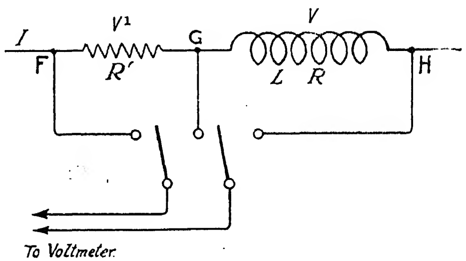


FIG. 88.

and V. Then

$$V' = R'I, \; V = ZI,$$

whence

$$Z = R'V/V'.$$

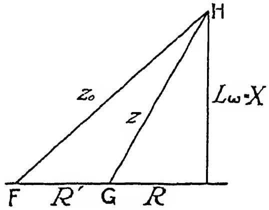


FIG. 89.

Now measure the voltage across FH and let this be V_0. If the augmented impedance be Z_0 we have

$$Z_0 = R'V/V_0.$$

The vector diagram is drawn in Fig. 89, the oscillatory current through the circuit being assumed as 1 ampere. From the diagram we see that

$$Z_0^2 = (R' + R)^2 + X^2$$

and

$$Z^2 = R^2 + X^2$$

whence

$$Z_0^2 - Z^2 = R'^2 + 2RR'.$$

Thus

$$R = (Z_0^2 - Z^2 - R'^2)/2R'$$

and

$$X = (Z^2 - R^2)^{\frac{1}{2}}.$$

The two switches sketched in Fig. 88 can easily be manipulated to put the voltmeter across R', Z and Z_0 in turn.

Resistance and Capacitance in Series.

52. The diagram for this case appears in Fig. 90 and is quickly sketched by remembering that the condenser reactance voltage is opposite in sense to that of inductive reactance, and has its vector always drawn lagging a right angle behind the ohmic vector. The resultant voltage OB is derived from the triangle, and has the equation

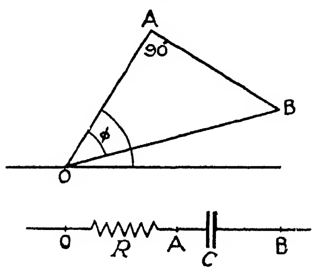


FIG. 90.

$$OB^2 = OA^2 + AB^2$$

or

$$E^2 = (RI)^2 + (I/C\omega)^2.$$

Let the square root of this be written $E = ZI$. Then

$$Z = (R^2 + 1/C^2\omega^2)^{\frac{1}{2}}$$

is the impedance of a resistance and a condenser in series. In this case it is the current that leads the voltage, and the diagram shows that the phase angle is given by

$$\tan \phi = -1/RC\omega,$$

where the minus indicates that the voltage is lagging behind the current. As in § 49 OA is the power component, AB the wattless component, of the voltage.

Inductance and Capacitance in Series. Resonance.

53. This paragraph introduces us to a very important part of our whole subject, a part to be studied more carefully, it might be said, than any other; for here we enter upon the theory of the electrical oscillator. In order to help to a clear understanding of the salient facts the mechanical analogy offered by a simple pendulum has been somewhat fully discussed in the section beginning with § 97, p. 162. Those readers who derive great assistance from mechanical analogies should read that section before reading the immediately succeeding paragraphs.

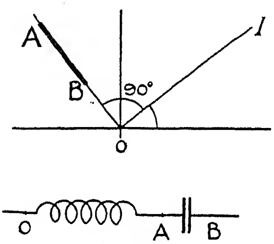


FIG. 91.

Through the coil and the condenser in Fig. 91 we suppose again the current I to run, producing the P.D. $L\omega I$ between the terminals of the coil and $(1/C\omega)I$ between the terminals of the condenser. These are perpendicular to the vector OI taken in any direction to represent the current, but the former leads and the latter lags. Thus OA $= L\omega I$, and AB $= I/C\omega$ and must be set off in the opposite direction from OA. Thus AB subtracts from OA, or the condenser tends to cancel the effect of the coil. The difference OB is of magnitude

$$(L\omega - 1/C\omega)I \equiv XI.$$

From the diagram it is clear that the terminal voltage across either the coil or condenser may be greater than the terminal voltage across the two, which may become very small as $1/C\omega$ gets nearer to $L\omega$ in size, vanishing in the limit when $L\omega = 1/C\omega$. This adjustment to numerical equality of the coil reactance and the condenser reactance, whether by altering L or C or ω, is called adjustment to "resonance," or "tuning." Near this adjustment only a very small voltage is needed to drive a large current through the assemblage. The condenser is said to have cancelled or annulled all the inductance. In this extreme case OB becomes

zero, which implies that an infinitesimal voltage of frequency $\omega/2\pi$ would drive a finite current through the pair. In other words, the current would rise to an indefinitely large value if the voltage were kept constant and if the source could furnish an indefinitely large current. From a different point of view we may say that a small voltage applied to the two in series gives rise to a magnified voltage across each.

Numerical Examples of an Electrical Oscillation.

54. Suppose the frequency of supply 50 ~, so that $\omega = 314$, and let $L = 2$ henrys. Then the reactance is 628 ohms. To annul this requires a negative reactance—that is, a condenser reactance—of 628 ohms. Let $1/C\omega = 628$ ohms then

$$C = 1/628 \times \omega \text{ farad} \doteqdot 5\mu\text{F}.$$

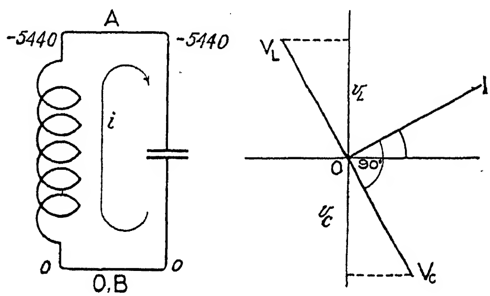


FIG. 92.

If L and C were connected to form a closed circuit as in Fig. 92, its natural frequency would be 50 ~, and if started by an electric blow it would oscillate at this frequency for a long time as the resistance is negligible. Let us suppose the current in the direction marked in Fig. 92 to be increasing at this moment, then the vector OI will be in the first quadrant. The potential drop v_L along L is indicated by the vector $OV_L = L\omega I$, and is positive and diminishing. The potential drop v_C across the condenser is indicated by aid of the vector $OV_C = I/C\omega$, and is seen to be negative and decreasing in numerical value as the cranks rotate. Suppose the maximum current to be $I = 10$ amperes and the time of the diagram 1/600 second after $t = 0$. Then at this instant

$$\omega t = 2\pi \,.\, 50/600 = 0{\cdot}523 \text{ radian}$$
$$= 360 \,.\, 50/600 = 30 \text{ degrees},$$

and therefore $\quad i = I \sin \omega t = 10 \times 0{\cdot}5 = 5$ amperes.

The vector OV_L is given by

$$V_L = L\omega I = 628 \times 10 = 6{,}280 \text{ volts}$$

and $\quad V_C = V_L = 6{,}280$ volts.

The instantaneous value of the potential drop in the inductance is

$$v_L = V_L \cos \omega t = 6{,}280 \times 0{\cdot}866 = 5{,}440 \text{ volts.}$$

Similarly $\qquad V_C = -5{,}440$ volts.

The potential drop across the condenser is negative, which means that the potential at the instant considered is greater at the plate from which the current is flowing than at the other plate. The numerals in Fig. 92 indicate the distribution of potential.

55. A radio frequency example may be treated in the same way. Taking $\omega = 10^6$ radians per second, which corresponds to a frequency of 159,000 per second, and taking $C = 1$ billifarad, we find the reactance is

$$-1/10^{-9} \times 10^6 = -1{,}000 \text{ ohms.}$$

For resonance the reactance of the coil must balance this, and therefore

$$L = 1{,}000/\omega = 10^{-3} \text{ henry}$$
$$= 1 \text{ millihenry.}$$

Resistance, Inductance and Capacitance in Series.

56. When the vectors for this case are drawn end to end, the appearance of the diagram depends greatly on the relative magnitudes of the positive and negative reactances. We may

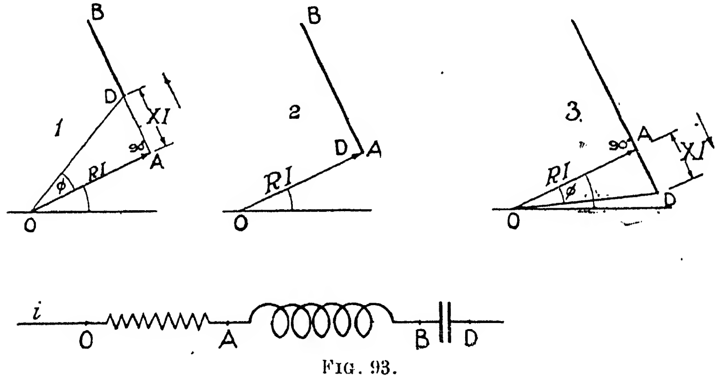


FIG. 93.

have the positive reactance predominant as in (1) of Fig. 93, or the negative predominant as in (3) of Fig. 93, with the dividing case where positive reactance is cancelling all the negative as in (2). Since the total reactance is of magnitude $X = L\omega - 1/C\omega$, then case (1) is that of X positive, (3) is that of X negative, and

(2) is the case $X = 0$. In all cases, we have from the triangles formed of RI, XI, and ZI,

$$Z^2 = R^2 + X^2 = R^2 + (L\omega - 1/C\omega)^2$$

$$\phi = \lrcorner X/R = \lrcorner (L\omega - 1/C\omega)/R.$$

Thus if inductance predominates ϕ is positive, and the voltage leads the current, and conversely if the condenser reactance is predominant.

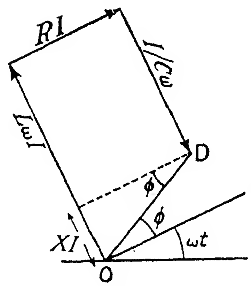


FIG. 94.

In Fig. 93 arrows have been marked on the vector diagrams. These are to show the directions taken as positive. In the diagram of connections the same direction must be taken as positive for the current and the voltage.

57. The three vectors representing the ohmic and the two reactive potential drops may be arranged end to end in several different ways. Taking the inductive case of Fig. 93, where the ohmic vector is set out first, we may, instead, set out $L\omega I$ first, the ohmic vector second, and the condenser vector last. We thus obtain Fig. 94. Here the resultant OD has the same magnitude and has the same direction relative to OI as would be obtained by the former way of adjoining the vectors.

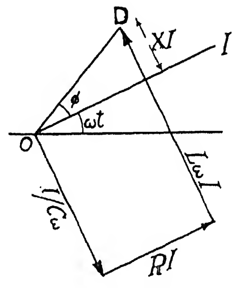


FIG. 95.

Another way of drawing the vector diagram arises on taking the condenser vector first and the coil vector last. We then obtain Fig. 95. The resultant OD has, of course, the same magnitude and direction as in the alternative diagrams.

Numerical Example.

58. Let it be required to find the potential drop between the points marked PQ in the circuit of Fig. 96, when a current of 10 amperes is flowing and where $\omega = 10^6$, $L = 0{\cdot}6$ mH, $C = 1{\cdot}4$ bF, $R = 450\ \Omega$; and let the resistance between P and B be $250\ \Omega$ and the inductance between B and Q $0{\cdot}4$ mH. The reactance of

L is 600 Ω, of C is 714 Ω, and therefore the potential drops in them are 6,000 V and 7,140 V. The ohmic drop is 4,500 V. Set these out to scale as in Fig. 96. On the vector AB cut off BP = 2,500 V and on BD cut off BQ = 4,000 V. Join PQ; it gives the voltage between P and Q and its phase relative to any other vector. The voltage is

$$V = \sqrt{(PB^2 + BQ^2)} = \sqrt{(2{,}500^2 + 4{,}000^2)} = 4{,}720 \text{ V}.$$

Now the phase angle between this voltage and the current is the angle QPB. This is given by

$$\tan\phi = QB/BP$$
$$= 4{,}000/2{,}500$$
$$= 1{\cdot}6$$

or $\phi = 58$ degrees.

By moving the tappings P and Q, the phase and voltage taken off may be shifted through a right angle, and, therefore, by aid of occasional changes of the capacitance can be placed at any angle with the applied voltage.

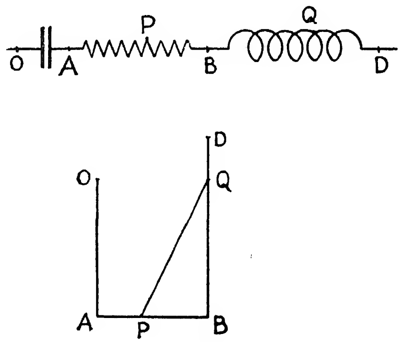


FIG. 96.

RESONANCE AND RESONANCE CURVES

59. Inspection of the diagrams or of the formulæ in the general case shows that the impedance is least when $X = 0$, and that then $Z = R$, $\phi = 0$. Thus if the positive and negative reactances annul, the assemblage behaves as if only the resistance were present, so far as regards the amplitude of the current flowing and its phase relative to the terminal voltage. In this state the assemblage is said to be " tuned to " or " in resonance with " the frequency of the source of emf, and the current is the greatest possible, for any adjustment of L or C, for a given voltage. If the voltage be kept of constant amplitude then an alteration of L or C from the resonance values will in general reduce the current and introduce a phase difference between current and voltage, either lag or lead. These changes in current and phase are represented in the radial Argand diagrams of Fig. 97 and in the

Cartesian diagrams of Fig. 98. In the former OA $= R$, to scale, and AP $= X$, the reactance at any particular adjustment; hence OP $= Z$. Let OB represent the current amplitude at the resonance adjustment, that is OB $= V/R$, and on it as diameter

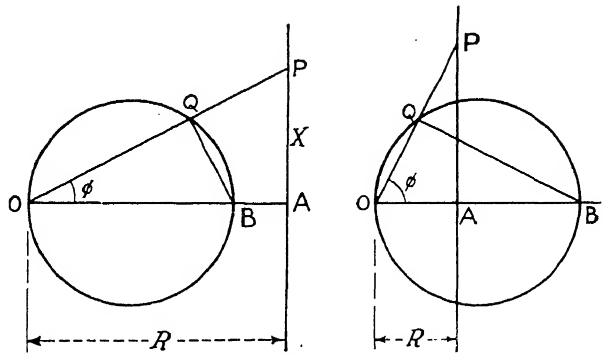


FIG. 97.

draw a circle cutting OP in Q. Join QB. In the two similar triangles OBQ, OPA we have

$$\begin{aligned} \text{OQ/OB} &= \text{OA/OP} \\ \therefore \text{OQ} &= (V/R) \times (R/Z) \\ &= V/Z \\ &= I. \end{aligned}$$

Again, from the Figure, $\tan \phi = X/R$.
Thus as X is varied by alteration of L or C or both, OQ gives the current amplitude and the phase angle marked ϕ. These results are transferred to Fig. 98 where the curve marked I is one form of resonance curve. The reactance X is positive or negative according as inductance or capacitance predominates.

Resonance when Frequency is Varied.

60. In the above discussion of the resonance curve it has been supposed that the oscillatory assemblage was traversed by a current of constant frequency and the reactance varied by variation of condenser or inductance. But the value of the reactance X may also be altered by changing only the frequency of the current, for ω is involved in the equation for X, namely,

$$X = L\omega - \frac{1}{C\omega}.$$

In the equation of the resonance curve

$$I = \frac{V}{\sqrt{(R^2 + X^2)}}$$

let L and C have fixed values and suppose ω to increase from zero; then X begins with a very large negative value, becomes zero, and passes through increasing positive values. Since X is squared, the negative values of X produce the same mathematical effect in the resonance equation as the positive values, and the resonance curve is the same shape as before, if values of X are used as abscissæ. If, on the other hand, values of frequency are used as abscissæ, the curve becomes unsymmetrical in shape and

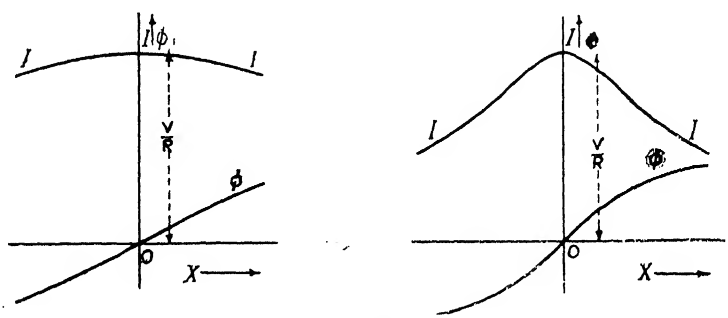


FIG. 98.

less simple geometrically. It is therefore better whenever possible to use values of X as abscissæ, though not essential.

61. An application of the resonance curve may be pointed out in this connection. Consider the state of things at a point on either sloping flank of the curve which corresponds to the value of I produced in the circuit at a definite frequency. If the frequency of the source of emf should alter while the voltage remains constant, then the representative point on the curve will move to the left and the current will fall, and *vice versâ*. Thus it is possible to use an ammeter with these circuits as a check on the constancy of the frequency of a supply.

62. The difference between " resonance " or " tuning " adjustment and the adjustment for " synchronism " should be carefully noted. This latter adjustment means that for which the natural

frequency is the same as that of the applied force. Now, it is shown in § 46 and more fully in books on damped wave telegraphy that the natural frequency of a circuit comprising series resistance is given by

$$p = \sqrt{\left(\frac{1}{LC} - \frac{R^2}{4L^2}\right)}$$

Using this the difference between resonance and synchronism can be exhibited on a vector diagram. In Fig. 99 the line AB was first marked off to represent R to scale (unit current being imagined), and a line drawn perpendicular to it at each end. From the centre M with radius MD $= \sqrt{L/C}$ the point D is determined. Then BD is equal to Lp as may easily be proved by the reader. Now bisect MB at N, join the bisection to D, and draw a perpendicular to ND through D to cut the left-hand reactive vector at O′. It can be shown that O′A $= 1/Cp$. The join O′D then expresses fully the voltage required across the terminals of the assemblage in order to force 1 ampere through it at the frequency $p/2\pi$. In other words O′D is the impedance of the assemblage at the natural frequency. If we now draw AO equal to MD we determine a length OA to represent the adjustment

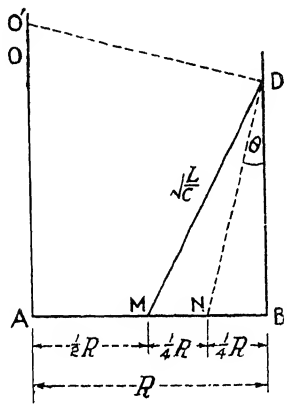


FIG. 99.

$$1/C\omega = L\omega,$$

that is to say, the resonance adjustment. It is obvious from the construction that the divergence between synchronism and resonance increases when R increases if the ratio L/C is unaltered. The phase difference is equal to the angle NDB and is expressible as

$$\begin{aligned}\tan\theta &= \text{NB/BD}\\ &= R/4Lp.\end{aligned}$$

Voltage Magnification by Resonance.

63. Let us take a series circuit not quite adjusted to resonance with a crank diagram such as that of Fig. 100. The emf applied

to the series assemblage is $E = ZI$. Let V be the voltage across the inductance coil. Then $V = L\omega I$. Thus we find

$$V/E = L\omega/Z.$$

Here $$Z = (R^2 + X^2)^{\frac{1}{2}}$$

and X is small because of nearness to resonance. If then R is greater than X we have approximately

$$\text{voltage magnification} = L\omega/R.$$

If R itself is small this fraction may be large.

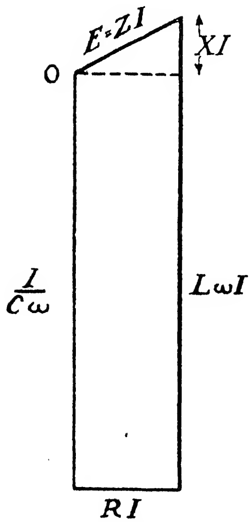


FIG. 100.

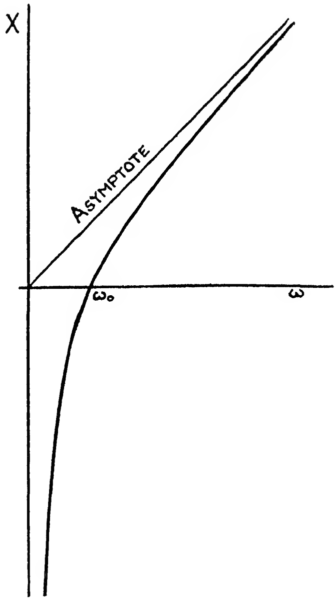


FIG. 101.

64. As a numerical example, let $L = 10^{-3}$ H, $C = 10^{-9}$ F, so that for resonance $\omega^2 = 1/(10^{-3} \times 10^{-9})$ or $\omega = 10^6$, and let the resistance be 1 Ω. Suppose a voltage of 100 volts to be applied. The current will be $100/1 = 100$ amperes. Therefore the amplitude of the terminal voltage of the inductance will be $L\omega I = 10^{-3} \times 10^6 \times 10^2 = 100{,}000$ V, and that across the condenser will be equal and opposite. It will be seen that it is only the existence of the resistance limits the rise of voltage and current to enormous values.

Variation of Reactance with Frequency.

65. The mode of variation of the reactance of the series circuit of Fig. 96 as frequency increases is exhibited by the curve of Fig. 101. Writing the equation

$$X = L\omega - 1/C\omega$$

in the form

$$\omega(\omega - X/L) = 1/LC = \omega_0^2,$$

where ω_0 is the resonance angular velocity, we see that the curve is a hyperbola with asymptotes $\omega = 0$ and $X = L\omega$ and that it cuts the axis of ω at the resonance value of ω, that is ω_0.

IMPEDANCES IN PARALLEL

66. The term impedance includes the extreme cases of pure resistance and pure reactance. Therefore among the possible parallel combinations we have to consider :—

(1) A pair of resistances.
(2) A pair of inductances.
(3) A pair of capacitances.
(4) Resistance in parallel with inductance.
(5) Resistance in parallel with capacitance.
(6) Inductance in parallel with capacitance.

More complicated cases follow in obvious arrangements, but the ones enumerated are of especial importance.

Resistances in Parallel.

67. Let two resistances R, R' have their terminals connected. Then the terminal voltages when they are traversed by a current will necessarily be equal and in the same phase. Let this common terminal voltage be v. Then calling the currents j, j' we have

$$j = v/R, j' = v/R'$$

and

$$i = j + j' = v\left(\frac{1}{R} + \frac{1}{R'}\right).$$

Thus

$$\frac{i}{v} = \frac{1}{R} + \frac{1}{R'} \text{ or } \frac{v}{i} = \frac{RR'}{R + R'}$$

but v/i is the effective resistance of the pair. Thus the resistance of two parallel resistances is their product divided by their sum. Again,

$$j = \frac{v}{R} = \frac{R'i}{R + R'} \text{ and } j' = \frac{v}{R'} = \frac{Ri}{R + R'}.$$

The current in one branch is the same fraction of the sum of the currents as the resistance of the other branch is of the sum of the resistances.

For several resistances in parallel, the same reasoning gives

$$\frac{i}{v} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} +, \text{etc.},$$

which is the easiest form of the equation and may be written $i =$ (sum of the conductances) v. The conductance of a coil is the reciprocal of its resistance and is measured in "mhos."

Inductances in Parallel, Capacitances in Parallel.

68. Precisely the same reasoning holds in this case as in the last, and we therefore have for the resultant reactance

$$\frac{1}{L_1D} + \frac{1}{L_2D} + \frac{1}{L_3D} + \ldots \text{ or } \frac{1}{L_1\omega} + \frac{1}{L_2\omega} + \frac{1}{L_2\omega} +$$

where the D in the denominators involves rotation through a right angle backward in the crank diagram.

Again we have for condensers

$$C_1D + C_2D + C_3D + \ldots \text{ or } (C_1 + C_2 + C_3 + \ldots)\omega$$

where the D warns us that the vector in the crank diagram is rotated forward by a right angle.

General Case of Impedances in Parallel.

69. The general case may now be taken algebraically, for it has been established in preceding pages that in oscillatory equations the resistance symbols may be generalised by writing the vector impedance $R + XD/\omega$ or (Z, ϕ) in place of R. Thus we get at once for $(Z, \phi) = R + XD/\omega$ and $(Z_1, \phi_1) = R_1 + X_1D/\omega$ connected in parallel

$$\frac{(Z, \phi)(Z_1, \phi_1)}{(Z, \phi) + (Z_1, \phi_1)} = \frac{(ZZ_1, \phi + \phi_1)}{(Z_2, \phi_2)}$$
$$= \left(\frac{ZZ_1}{Z_2}, \phi + \phi_1 - \phi_2\right)$$

by § 42. Written out in full we have

$$Z = (R^2 + X^2)^{\frac{1}{2}}, Z_1 = (R_1{}^2 + X_1{}^2)^{\frac{1}{2}},$$
$$Z_2 = \{(R + R_1)^2 + (X + X_1)^2\}^{\frac{1}{2}}, \phi = \angle X/R,$$
$$\phi_1 = \angle\frac{X_1}{R_1}, \phi_2 = \angle\frac{(X + X_1)}{(R + R_1)}.$$

The rules for a pair of parallel impedances may be put in the same terms as those given in § 67 for resistances, merely with the word resistance changed to vector impedance.

70. For several impedances in parallel, the same reasoning as was used for several resistances leads to the formula

$$\frac{1}{(Z, \phi)} = \frac{1}{(Z_1, \phi_1)} + \frac{1}{(Z_2, \phi_2)} + \frac{1}{(Z_3, \phi_3)} + \dots$$

where (Z, ϕ) is the whole impedance, $(Z_1, \phi_1) \equiv R_1 + X_1 D/\omega$, etc.

The reciprocal of the impedance of a coil is called its "admittance" and is measured in "mhos."

Particular Cases.

71. Although the general formulæ and rules just given cover all cases, it is instructive to examine certain of the most commonly occurring cases, with especial attention to the crank diagrams. Current vectors are appropriate for parallel arrangements just as voltage vectors are characteristic of series arrangements. But as it is nearly always best ultimately to work with crank diagrams of voltages, current diagrams will not be much used in these pages.

Resistance and Inductance in Parallel.

72. We shall use a "circuital" or "cyclic" notation for the currents. In Fig. 102 the arrows indicate the positive directions of the currents, not the directions at any particular instant. The current in L at any instant is $I_2 \sin \omega t$, that in R at the same instant is $I_1 \sin(\omega t + \phi) + I_2 \sin \omega t$. In most of our work it will be found convenient to take the positive directions in the same sense in that member which is traversed by the two currents. It does not matter which is made the common member, as will be seen in a later paragraph.

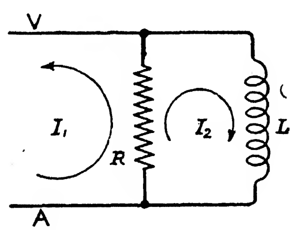


FIG. 102.

Now in any circuit in which there is no source of emf the sum of the potential drops in all the parts reckoned in the same

sense round the circuit must be zero; or if there is a source of emf the sum must be equal to the applied emf. This electrical fact, taken to the crank diagram, means that the geometrical sum of the voltage cranks must be zero (or equal to the applied emf, if any). In other words, the voltage vectors belonging to a circuit must form a closed figure. This principle is sufficient to enable all of them to be determined when the merely electrical data are complete.

Thus in circuit 2 the potential drop in L has a crank of length $L\omega I_2$ and it leads the current by a right angle. That in R has a crank RI_2 due to current I_2 and a crank of length RI_1 due to current I_1. The phase between these ohmic P.D.'s is not known yet, but becomes known on drawing the triangle of vectors. For mark off $L\omega I_2$ to scale in any direction as at VA in Fig. 103. If $L = 1$ mH and $\lambda = 1{,}885$ metres we shall have $L\omega = 1{,}000$ ohms, so it will be convenient to take $I_2 = 1/100$ ampere, and to use a scale $1'' = 1$ volt. Then draw RI_2 on the same scale as a crank AB a right angle behind. There is only one way of closing the triangle, and this third line must represent the third voltage in circuit 2, namely, RI_1. As for circuit 1, whatever there may be on the left of the terminals AV, the potential drop between A and V is obtained by adding cranks RI_2 and RI_1. This is done in Fig. 104. Of course there is no need to draw these diagrams separately; for when they are compared it is seen that V is identical with $L\omega I_2$ reversed, as is immediately obvious from an inspection of the circuit. From the diagram we can deduce the impedance of the assemblage

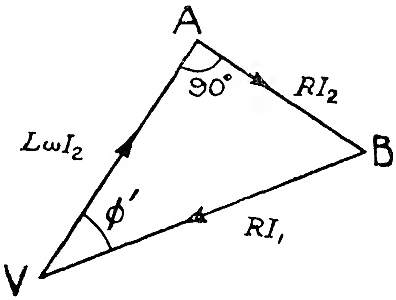


FIG. 103.

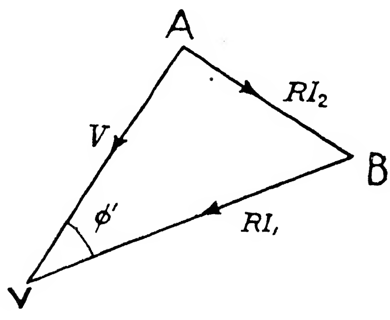


FIG. 104.

by noticing that, by definition, $V = Z'I_1$, or that the impedance is the ratio of V to I_1. From this we see that

$$Z' = R\frac{\mathrm{AV}}{\mathrm{BV}} = \frac{RL\omega}{(R^2 + L^2\omega^2)^{\frac{1}{2}}}.$$

If $R >> L\omega$, that is if the power factor is small

$$Z' \doteqdot L\omega(1 - \tfrac{1}{2}L^2\omega^2/R^2)$$

and if $R << L\omega$ or the power factor large

$$Z' \doteqdot R(1 - \tfrac{1}{2}R^2/L^2\omega^2).$$

The phase angle of the assemblage is

$$\begin{aligned} \phi' &= \angle \mathrm{AVB} \\ &= \lrcorner R/L\omega. \end{aligned}$$

These results can also be obtained by the method of the resistance operator.

73. The power factor is $\cos\phi'$, that is AV/VB, which equals $L\omega/(R^2 + L^2\omega^2)^{\frac{1}{2}}$. Now the work done on the circuit per second is, by § 30,

$$\begin{aligned} P &= VI_1 \cos\phi' \\ &= Z' \cos\phi' \,.\, I_1^2. \end{aligned}$$

This work may be imagined to be spent in an "equivalent" or "image" resistance, that is a resistance capable of representing the circuit so far as energy-absorbing is concerned, and then we should have

$$P = R'I_1^2.$$

Thus

$$\begin{aligned} R' &= Z' \cos\phi' \\ &= \frac{RL^2\omega^2}{R^2 + L^2\omega^2}. \end{aligned}$$

When $R >> L\omega$ this equation becomes

$$R' \doteqdot \frac{L^2\omega^2}{R}\left(1 - \frac{L^2\omega^2}{R^2}\right)$$

or the image resistance is proportional to the square of the frequency approximately. When, on the other hand, $R << L\omega$ we have

$$R' \doteqdot R(1 - R^2/L^2\omega^2)$$

and is almost independent of frequency. The former case arises in closely wound coils of wire when the insulation between turns and layers is so bad that it acts as a shunt resistance.

Numerical Examples.

74. An inductance of 1 mH in parallel with a crystal detector of resistance 10,000 ohms at wavelength 1,885 metres has a phase angle whose tangent is $10{,}000 \div 1{,}000 = 10$. This angle is nearly a right angle, and therefore the power factor is low. Thus

$$Z' \doteqdot 1{,}000(1 - \tfrac{1}{2} \cdot \tfrac{1}{100}) = 995 \text{ ohms}.$$

The effective or image resistance is

$$R' = \frac{10^6}{10^4}\left(1 - \frac{10^6}{10^8}\right) = 99 \text{ ohms}.$$

Note that although the reactance $L\omega$ is 1,000 ohms, it gives a different effective resistance from that given by 1,000 ohms of non-inductive resistance in parallel with the detector, which would give 909 ohms.

Again, if a choking coil of inductance 10 H and of negligible resistance is connected in parallel with a filament lamp of resistance 150 ohms and put on a supply of 50 $\sim$ current at 100 volts the ratio of the current through the lamp to the current through the choking coil is inversely as their reactances, and is therefore LD to R. This is $10 \times 2\pi \times 50$ to 150, which equals 220 to 21 or more than 10 to 1 in magnitude. The D means that the current through the resistance is leading the current through the choking coil by a quarter period. The ratio of the amplitudes being more than 10 to 1, the bulk of the alternating current will pass through the lamp, and it will glow nearly as brilliantly as if alone. On the other hand, if a steady current were passed through the combination practically none of the current would go through the lamp, because by hypothesis the resistance of the coil is infinitesimal compared with that of the lamp. The working out of the impedance and power factor is left to the reader.

Equivalence of Different Crank Diagrams.

75. In arranging the circuital currents I_1 and I_2 in Fig. 102 there was a certain arbitrariness in choosing that R and not L should be traversed by both currents. The question arises: Would different results be obtained if the circuits had been drawn as in Fig. 105 wherein the inductance has been chosen as the common part of the circuital currents? It is plain that the sketches are electrically identical and that the current passing

the terminals AV will again be I_1, if the P.D. is V, but the new symbol J_2 must be assigned to the new circuital current.

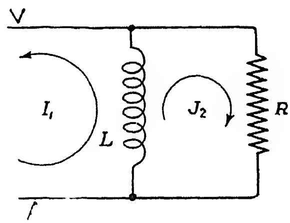


FIG. 105.

The triangle of vectors for circuit 2 is made up RJ_2, $L\omega J_2$ and $L\omega I_1$. This is indicated in Fig. 106. Taking RJ_2 in any direction VA, $L\omega J_2$ can be drawn at once because it must be perpendicular to the ohmic drop. Having these two sides in position, the third side B′V must represent the drop $L\omega I_1$. This last line defines the direction and size of the current I_1, which must be a right angle behind the drop vector. It is drawn in and marked I_1. Now the drop between the terminals A and V may be reckoned as the geometrical sum of cranks $L\omega J_2$ and $L\omega I_1$, that is of AB′ and B′V, which is AV. It may also be reckoned as the terminal voltage of R and is then RJ_2 in opposite sense to the positive direction of J_2. The phase angle of the assemblage is marked ϕ'; it is also the angle AB′V. Its tangent is $R/L\omega$. The resultant impedance is equal to the ratio V/I_1, which is the ratio VA $\div$ (VB′$/L\omega$), which leads to

$$Z = \frac{RL\omega}{(R^2 + L^2\omega^2)^{\frac{1}{2}}}.$$

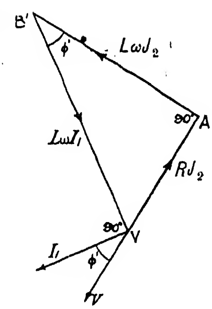


FIG. 106.

Similarly all the results that follow from Fig. 103 come also from Fig. 106. Thus the arbitrary choice of symbols for the circuital currents is, as was to be expected, of no significance.

Resistance and Capacitance in Parallel.

76. As it is indifferent whether we take the condenser or the resistance in Fig. 107 as the part common to the two circuital currents I_1 and I_2, we take the condenser in the present problem for variety. Also, instead of starting from I_2 we shall start by assuming I_1 given. Taking its crank in any direction, the crank for the component drop due to that current in C will be perpendicular to and behind the current. Set it off to scale at AB in

Fig. 108. From B try to set off BV to represent the other component of the drop across the condenser, namely, that due to the current I_2. Perhaps the best way of finding the correct direction for BV is to use the fact that the two vectors due to I_2 must form a right angled triangle with AB as hypotenuse and with the angle ABV equal to $\lrcorner$ AV/VB = $\lrcorner\, RC\omega$. When R, C and ω are known this angle is easily calculated and marked off at B. Then a perpendicular from A cuts the boundary of the angle at V, thus determining I_2. It is clear from the circuit diagram that the drop between the terminals is $-RI_2 = V$, say. That is to say it is AV.

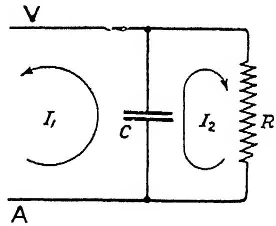


FIG. 107.

The phase angle of the shunted condenser is the angle between V and I_1, and is marked ϕ' in Fig. 108. It is equal to the angle ABV. From the diagram we have for the shunted condenser :—

$$\text{Phase angle} \qquad \phi' = \lrcorner\, RC\omega$$

$$\text{Power factor} \qquad \cos\phi' = \frac{1}{(1 + R^2C^2\omega^2)^{\frac{1}{2}}}$$

$$\begin{aligned}\text{Effective impedance } Z' &= V/I_1 \\ &= \text{AV} \div \text{AB} \,.\, C\omega \\ &= R/(1 + R^2C^2\omega^2)^{\frac{1}{2}}\end{aligned}$$

$$\text{Effective resistance } R' = R/(1 + R^2C^2\omega^2).$$

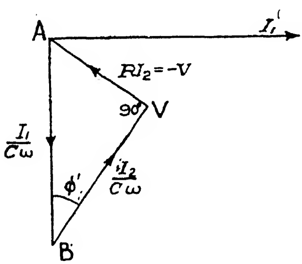


FIG. 108.

Also the ratio of the circuital currents, namely,

$$I_2 : I_1, \text{ is VB : AB} = \cos\phi'.$$

77. A condenser is sometimes used for shunting a hot wire ammeter in radio-frequency laboratories because of the difficulty of procuring a really non-inductive resistance coil for the purpose. Let us suppose an ammeter of 10 ohms resistance is to be used for measuring currents 10 times greater than it can measure directly. What is the magnitude of the condenser shunt needed? Regarding R in the

above formulæ as belonging to the hot wire ammeter, we must have $I_2/I_1 = 1/10$. By the last paragraph we have

$$0{\cdot}1 = (1 + R^2C^2\omega^2)^{-\frac{1}{2}}.$$

Hence $$C^2 = 99/R^2\omega^2 = 0{\cdot}99/\omega^2$$

or $$C\omega = 0{\cdot}995 \text{ mho}.$$

The condenser required is therefore different at every frequency. Suppose the frequency corresponds to $\omega = 10^6$. Then the condenser must have capacitance $0{\cdot}995\ \mu$F.

Any Reactance in Parallel with Resistance.

78. The circuit of Fig. 109 is arranged with the resistance as the common conductor of the two circuital currents. Assuming I_1 to be given we first draw vector AB (Fig. 110) $= RI_1$ to scale and then try to draw the triangle ABV with sides BV and VA in

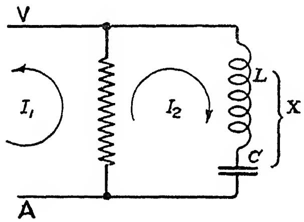


FIG. 109.

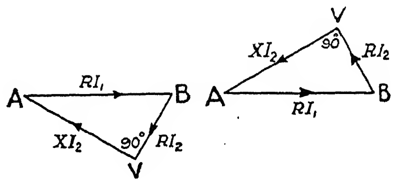


FIG. 110.

proportion to the given value of R and X. This is easily done by calculating the angle ABV $= \angle X/R$. In the left-hand crank diagram of Fig. 110 X is negative, in the right-hand diagram it is positive—that is, on the left hand $1/C\omega > L\omega$, on the right $L\omega > 1/C\omega$.

Inductance and Capacitance in Parallel.

79. The amount of space devoted to mathematical treatment will be reduced by taking a very general case first, namely, that in which the coil and the condenser have each a resistance in series. Fig. 111 explains the symbols adopted. Either one or both of the resistances may be put zero to obtain particular cases and thus a great number of problems occurring in wireless telegraph practice can be covered by one set of formulæ.

The separate resistance operators are $LD + S$ and $R + 1/CD$. The shunt rule of § 67 gives the resultant operator as

$$\frac{(S + LD)(R + 1/CD)}{S + LD + R + 1/CD}.$$

This may be multiplied out and $-\omega^2$ put for D^2 as in § 37; but it is sufficient for our present purpose to take $R + LD$ as equivalent to multiplication by Z_1 and rotation of a crank through the angle ϕ_1, that is equivalent to (Z_1, ϕ_1); $S + 1/CD$

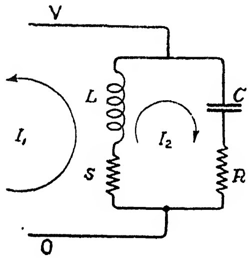


FIG. 111.

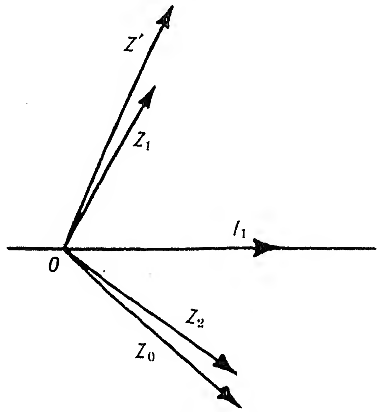


FIG. 112.

equivalent to (Z_2, ϕ_2), and $R + S + LD + 1/CD$ equivalent to (Z_0, ϕ_0). The operator then becomes

$$(Z_1, \phi_1)\,(Z_2, \phi_2)/(Z_0, \phi_0)$$

which $$\doteq (Z_1 Z_2/Z_0,\ \phi_1 + \phi_2 - \phi_0)$$

where Z_0 and ϕ_0 are in fact the impedance and the phase angle of the assemblage regarded as a closed isolated circuit.

The equivalent impedance of the two parallel branches is thus $Z' = Z_1 Z_2/Z_0$, and the phase angle of the assemblage is $\phi' = \phi_1 + \phi_2 - \phi_0$.

80. As an example let $\omega = 10^6$, $S = 100\,\Omega$,

$$L = 0{\cdot}2\,\text{mH},\ R = 120\,\Omega,\ C = 2{\cdot}5\,\text{bF}$$

then
$$Z_1 = (S^2 + L^2\omega^2)^{\frac{1}{2}} = 224\,\Omega$$
$$Z_2 = (R^2 + 1/C^2\omega^2)^{\frac{1}{2}} = 418\,\Omega$$
$$Z_0 = \{(R + S)^2 + (L\omega - 1/C\omega)^2\}^{\frac{1}{2}} = 297\,\Omega$$
$$\phi_1 = \rfloor L\omega/S = \rfloor 2 = 63^\circ\,28'$$
$$\phi_2 = \rfloor -1/RC\omega = -\rfloor 400/120 = -36^\circ\,51'$$
$$\phi_0 = \rfloor X/(R + S) = \rfloor -200/220 = -42^\circ\,18'$$

Hence
$$Z' = 224 \times 418 \div 297 = 315\,\Omega$$
$$\phi' = 63^\circ\,28' - 36^\circ\,51' + 42^\circ\,18' = 68^\circ\,55'.$$

These impedances are exhibited in Fig. 112, or, if we wish so to regard the vectors, they are the terminal voltage vectors when unit current represented by OI passes through the respective impedances. The currents are easily calculated when the voltage V is given, for $I_1 = V/(Z', \phi')$ and $I_2 = V/(Z_2, \phi_2)$. Let $V =$ 1,000 volts, then $I_1 = 2{\cdot}96$ A, and is 68° 55′ behind the voltage, and $I_2 = 2{\cdot}39$ A and is 36° 51′ ahead of the voltage.

We have also $J = V/(Z_1, \phi_1) = 4{\cdot}46$ A, lagging 63° 28′ behind the voltage, where J is the amplitude of the resultant current through the inductive branch, and is not the mere sum of I_1 and I_2.

Crank Diagram for Circuit of Fig. 111.

81. Assuming any direction for I_1, we have in Fig. 113 the potential drops OA and AB for that current through S and L. The current I_2 gives inductive, ohmic and condenser drops at BD, DV, VF and FO as marked in the diagram. The resultant of the ohmic and inductive drops in S and L due to both I_1 and I_2 is OV.

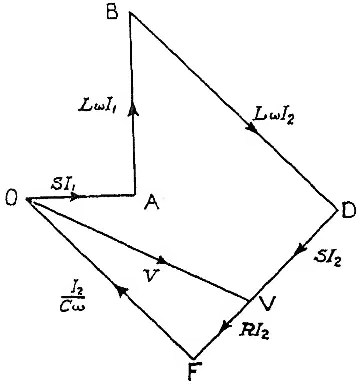


FIG. 113.

A Frequency Measurer.

82. From the above definition of J and I_2 and the rule of § 67 we obtain

$$\frac{I_2^2}{J^2} = \frac{S^2 + L^2\omega^2}{R^2 + 1/C^2\omega^2} = \gamma^2$$

where γ is a symbol denoting the ratio of the currents. Let us suppose $S = R$ and also that $L/C = R^2$. Then the above equation, which is a quadratic for ω^2, simplifies greatly, and we obtain

$$\omega^2 = \frac{\gamma^2}{LC}.$$

Now suppose that R and S are the resistances of two equal hot wire ammeters, say $R = S = 100$ ohms, and let L be 1 millihenry. Then $C = L/R^2 = 10^{-3}/10^4 = 10^{-7}$ farad $= 1/10$ microfarad, thus $\omega = \gamma \, . \, 10^5.$

Therefore with a pair of ammeters and appropriate inductance and capacitance simultaneous readings of the two instruments give by their ratio the frequency of the current passed through the assemblage. An instrument utilising this principle has been described by G. Ferrié.

Equivalent Resistance and Reactance of Parallel Connections.

83. We shall obtain formulæ for the general case of the circuit of Fig. 114 where the nature of the reactance is not specified. In drawing the crank diagram we shall suppose X and X_2 positive, but the algebraic results obtained will hold good when either

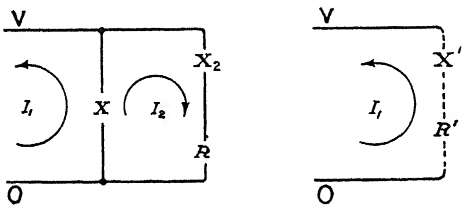


FIG. 114.

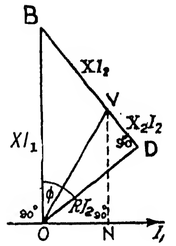


FIG. 115.

or both of these reactances are negative. The crank diagram is given in Fig. 115. It is evident, since OV represents the emf necessary to force the current I_1 through the compound circuit, that $ON = R'I_1$, $NV = X'I_1$, $OV = Z'I_1$, and $\angle NOV = \phi'$, where R', X', Z' and ϕ' are the equivalent resistance, reactance, impedance and phase angle of the given compound circuit.

84. A very important quantity, not included in the above enumeration, is the ratio of the cyclic current I_2 in the closed circuit to the current I_1 supplied to the assemblage. This is easily expressed in terms of the given electrical constants by writing down the value of the vector OB in two ways, namely, as XI_1 and as $(OD^2 + DB^2)^{\frac{1}{2}}$.

Thus
$$XI_1 = ZI_2$$
or
$$I_2/I_1 = X/Z.$$
Now ON is the projection of BV on the I_1 vector,

therefore
$$R'I_1 = XI_2 \cos\phi = XI_2R/Z$$
or
$$R' = RX^2/Z^2.$$
Here
$$Z^2 = (X + X_2)^2 + R^2.$$
Again NV is OB less the projection of BV on OB. Therefore

$$X'I_1 = XI_1 - XI_2 \sin\phi$$

or
$$X' = X\left\{1 - \frac{I_2}{I_1}\frac{X + X_2}{Z}\right\}$$
$$= X\left\{1 - \frac{X(X + X_2)}{Z^2}\right\}$$
$$= \{R^2 + X_2(X + X_2)\} X/Z^2.$$

The vector OV can be expressed in three useful ways :—

$$\text{OV} = E = Z'I_1 = Z_2I_2.$$

Hence
$$Z' = Z_2X/Z.$$

From this it follows that the power factor of the assemblage,

$$\cos\phi' = R'/Z' = RX/ZZ_2.$$

These formulæ for R', X', Z' and ϕ' convey no more information than is conveyed by the lettered triangle ONV of Fig. 115. In practice it is usually easier and always more instructive to draw the crank diagram than to use the formulæ. On the drawing board a fairly large scale can be adopted and then all the necessary accuracy is attained.

85. The results are, however, so important that it is worth while obtaining them in another way which will illustrate the operator method. We apply the rule for parallel resistances given in § 67. The operator corresponding to branch X is XD/ω, that to the other branch $R + X_2D/\omega$. The product of these is

$$RXD/\omega - XX_2,$$

remembering that $D^2 = -\omega^2$.

The sum of the operators is

$$R + (X + X_2)D/\omega.$$

Now the resultant resistance operator is equal to the above product divided by the sum, that is

$$\frac{RXD/\omega - XX_2}{R + (X + X_2)D/\omega}.$$

Multiply numerator and denominator by

$$R - (X + X_2)D/\omega$$

and substitute $-\omega^2$ for D^2 in the result.

We obtain as the new denominator

$$R^2 + (X + X_2)^2 \text{ or } Z^2.$$

For the numerator we obtain

$$RX^2 + \{R^2 + XX_2(X + X_2)\} D/\omega.$$

Equating the fraction to $R' + X'D/\omega$ we obtain finally

$$R' = RX^2/Z^2$$
$$X' = \{R^2 + X_2(X + X_2)\}X/Z^2.$$

These agree with the formulæ obtained from the crank diagram and lead to the equations for Z' and ϕ'.

Condition for Non-reactive Resultant.

86. In various applications of oscillatory circuits a non-reactive arrangement is sought. In the case of circuits of the type of Fig. 114 we derive information about the necessary relations between the parts of the assemblage by making NV zero either in its diagram or in the formula. Taking the latter first we merely write

$$R^2 + X_2(X + X_2) = 0$$

whence $$XX_2 = -(R^2 + X_2^2).$$

Since the right-hand side is always negative for all possible values of R and X_2 we see that X and X_2 must be chosen with opposite signs for the assemblage to be non-reactive; that is to say, either X or X_2 must be mainly capacitance, but not both. The last equation allows of numerical adjustment being made when R is given. Turning to the crank diagram we see that when NV is zero we may have a diagram of the kind shown in Fig. 116 where X_2I_2 is negative and subtracts from BV $= XI_2$, or we may have a diagram in which X is negative, and OB is drawn down the page, which can easily be constructed by the reader to make $\phi' = 0$.

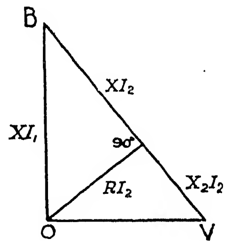


FIG. 116.

87. When the above adjustment is made we can deduce the current magnification of these arrangements—that is the ratio I_2/I_1—by equating the cosine of $\angle$OBV to the sine of $\angle$OVB. Thus

$$I_2/I_1 = -Z_2/R$$

where $$Z_2^2 = R^2 + X_2^2.$$

By expressing OV as $R'I$, and also as Z_2I_2 we obtain

$$R' = Z_2I_2/I_1 = Z_2^2/R.$$

The value of I_1 when the applied emf is E is

$$I_1 = E/R' = ER/Z_2^2.$$

Note that, as always, $R'I_1^2 = RI_2^2$.

Numerical Example.

88. Let R, X_2 in Fig. 114 be a coil of resistance 100 ohms and inductance one millihenry, and let the problem be to put a condenser in parallel to annul the inductance at $\omega = 10^6$.

Then by § 86

$$X \text{ or } -1/C\omega = -(10^4 + 10^6) \div 10^3$$
$$= -1{,}010 \text{ ohms}$$

and therefore $$C = (1/1{,}010)\mu\text{F}.$$

Also § 87 $$I_2/I_1 = -Z_2/R$$
$$= -(10^4 + 10^6)^{\frac{1}{2}} \div 10^2$$
$$= -\sqrt{101} \doteqdot -10$$

And therefore $$Z' \text{ or } R' = R(I_2/I_1)^2$$
$$= 10{,}100 \text{ ohms}.$$

We see from this example that the current running in the closed circuit is about 10 times that in the main and that the relative directions of the currents are opposite from those marked in the Figure. We see also that the closed circuit acts as a very large ohmic resistance placed in the main circuit.

Balance Wheel Circuits.

89. When a closed circuit is connected into a main circuit so that the two branches into which it becomes divided contain

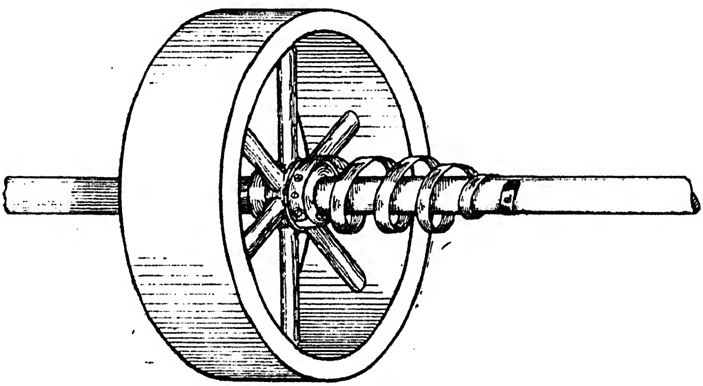


FIG. 117.
Loose Wheel on Smooth Shaft.

reactances of opposite sign it is said to be a "flywheel" or "balance wheel" circuit, provided that its parts are adjusted in size to approximately the values given in § 86. It behaves

somewhat like a loose flywheel might behave on a smooth horizontal shaft if the flywheel were coupled to the shaft by a spring like the hair-spring of the balance wheel of a watch, and then the shaft rotated to and fro about its axis by an alternating torque (Fig. 117). It is easy to see that a small alternating torque applied to the shaft would be transmitted to the wheel by the spring and would gradually build up an alternating rotatory motion of the wheel; and this would become the greater the smaller the frictional losses. In particular if the natural period of the balance wheel when the shaft was held fixed were the same as that of the applied alternating torque with which the shaft is to be moved, the motion built up might become very great relative to the motion of the shaft. To translate this to electrical terms we note that the rotatory motion of the shaft corresponds to the main current I_1 in Fig. 111 and Fig. 114, and that the rotatory motion of the flywheel corresponds to the current in the closed circuit, and we see that the current internal to the flywheel circuit may become much greater than the current in the main circuit when a certain relationship is approximately fulfilled. Obviously the inertia of the wheel corresponds to the positive reactance of one side of the electrical circuit and the springiness of the coupling between wheel and shaft corresponds to the negative reactance on the other side of the circuit.

90. To find the precise adjustment of X at which the ratio I_2/I_1 is a maximum we notice from § 84 that this ratio equals X/Z; whence

$$\left(\frac{I_1}{I_2}\right)^2 = \frac{R^2 + (X + X_2)^2}{X^2}$$

and this must be made a minimum. The ratio is

$$\frac{Z_2^2}{X^2} + \frac{2X_2}{X} + 1$$

which is of quadratic form and can be written

$$\left(\frac{Z_2}{X} + \frac{X_2}{Z_2}\right)^2 + 1 - \frac{X_2^2}{Z_2^2}.$$

This is clearly a minimum when the squared term is zero, that is when

$$Z_2/X = -X_2/Z_2$$

or

$$X = -Z_2^2/X_2.$$

When X has this value the ratio I_2/I_1 becomes

$$I_2/I_1 = -Z_2/R.$$

The minus shows that the current I_2 is in the opposite relative direction from that marked in Fig. 114.

91. If the value of X just arrived at be substituted in the general value of X' found from the crank diagram of Fig. 115 we obtain $X' = 0$. That is to say, we find we have arrived again at the adjustment discussed in § 86—the adjustment of X that gave the non-reactive condition.

92. When a series arrangement of resistance and reactance, such as that of Fig. 96, is connected into a circuit and adjusted to pass maximum current we find that the total reactance in circuit must be made zero and the resonance adjustment effected. In the corresponding parallel arrangement (Fig. 118) which was also discussed in § 90 we have, in contrast, to make

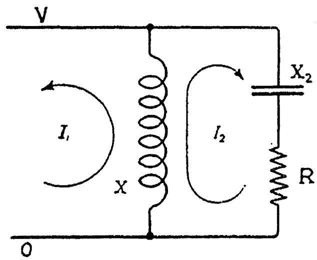


FIG. 118.

$$X = -Z_2^2/X_2$$
$$= -(R^2/X_2 + X_2)$$

or $X + X_2 = -R^2/X_2$.

The total reactance is therefore not to be made zero unless R is zero. If, however, we seek in the series arrangement the adjustment that gives the maximum P.D. across, say, the inductance when the inductance is varied and when the emf across the whole reactance and resistance is kept constant, we find that exactly the same mathematical relations must subsist as for the maximum current problem in the parallel connection just discussed. Moreover, the magnification of voltage effected in the series arrangement by the method of § 63 is equal to the current magnification obtained by the parallel connection of § 87 with the same relative adjustment between the two portions of the circuit.

As a rule in radio circuits the resistance is small compared with the reactance, and therefore the adjustments above described are approximately the true resonance adjustments.

Use of Crank Diagrams in following Changes in the Electrical Data

Reactance Variable.

93. Let us take the circuit of Fig. 109, where a resistance and a reactance are connected in parallel, and imagine the reactance X to be variable at will, while R and I_1 are kept constant. It is

evident that throughout all the changes the angle BVA in Fig. 110 must be a right angle because the ohmic drop BV and the reactive drop VA are due to the same current I_2 and are therefore in quadrature. Therefore as X varies the point V will describe a portion of a circle on OB as diameter. This is indicated in Fig. 119. For the position of V marked there the value of the reactance is positive and is about three times the value of R; but it started from an infinite value at B. As X decreases the chord OV decreases and BV increases. Since R is constant the last statement proves that I_2 is increasing though V (= OV) is falling. When $X = R$, V is at the point of the circle furthest from the diameter and the phase angle is 45 degrees. At this stage I_2 is 0·707 I_1, as is seen on comparing BV with BO. Still diminishing X we arrive at the zero value when V is at O. At this stage $I_2 = I_1$, and as these currents are opposite in phase there will be no current through R. In fact V is zero, that is to say, the voltage required to drive the current $I_2 = I_1$ through the reactance is infinitesimal.

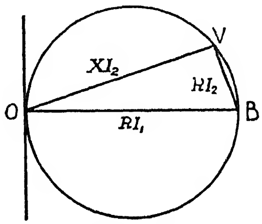


FIG. 119.

After this, if we still diminish X it becomes negative and the point V goes to the other side of the diameter OB. The voltage necessary to pass I_1 through the assemblage increases as V approaches B in this semicircle, till at last X is infinite, $I_2 = 0$, all the input current I_1 passes through R, and the emf is $V = RI_1$.

94. It is seen from this that the circle might be graduated with values of X to indicate the corresponding position for V. The easiest way of making the graduations is to draw through O a line perpendicular to OB and to set off along it lengths representing particular values of X to the same scale as OB represents R; the joins of the ends of the lengths to the point B will cut the circle in the points V corresponding to the values of X marked off.

The changes in X, which equals $L\omega - 1/C\omega$, may be due to changes in L or in C or in ω; it is quite indifferent for the purposes of constructing the diagram. In fact simultaneous changes in the above variables may be followed by aid of a little arithmetic and subsequent transference to the diagram.

95. It frequently happens that the reactance consists of only

inductance or only capacitance; that is to say $X = L\omega$ or $X = 1/C\omega$. The above described method of drawing and reading the circle locus is of course applicable to these special cases. When $X = L\omega$ the point V is always on the upper semicircle, when $X = 1/C\omega$ the point V is on the lower. This is a reminder of the fact that, mathematically speaking, an inductance may be regarded as a negative capacitance, and conversely.

Resistance Variable.

96. A brief inspection of Fig. 119 convinces one that this diagram is not a convenient one for following the results of changing the value of the shunt resistance; for two sides of the

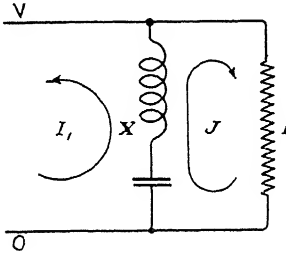


FIG. 120.

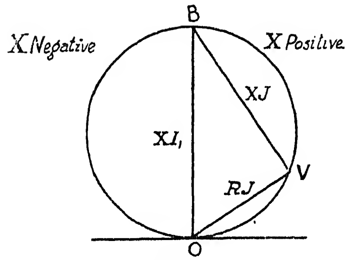


FIG. 121.—Reactance Constant.

triangle vary simultaneously because R varies, and the third side alters because I_2 does. With no part of the triangle fixed in size there will be difficulty in utilising it. But as has already been pointed out, the crank diagram can be drawn in another way that will introduce the variable quantity R into only one side of the triangle. The alternative method is given in Figs. 120 and 121. In the position of V here shown the reactance is positive; when it is negative the point V will be at the other side of the diameter.

A MECHANICAL ANALOGY

97. The preceding discussion of the oscillations of a simple electrical circuit will be helped by a review of the fundamental facts of a simple mechanical vibrator. In Fig. 122 a vice is represented holding a flat springy piece of steel, with its length horizontal, and on this strip is a movable bob that can be clamped in any position. It is evident without explanation that if the end of the spring be pulled horizontally to one side and released,

the bob and spring will perform oscillations in a horizontal plane, which will gradually decay, but will be of practically the same periodic time throughout their life.

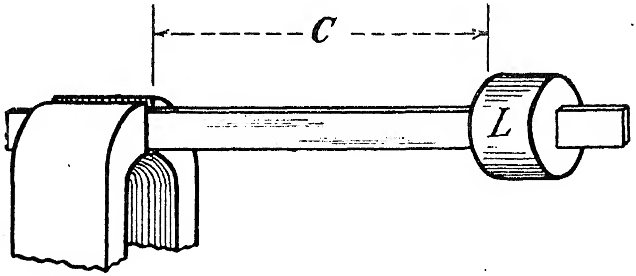


FIG. 122.

Let us analyse the motion. Suppose that a snapshot plan of the vibrator is represented by Fig. 123, the mass being on its outward journey from its middle position. At first we shall suppose there is no friction so that the vibrations would go on for a very long time and each to-and-fro motion would be exactly like its predecessor in every respect. At the instant depicted the spring is being bent and is elastically opposing the outward motion. It is therefore gradually reducing the speed and momentum of the bob, and according to a very fundamental law of mechanics the rate of reduction of the momentum is equal to the elastic force exerted by the spring at the instant. The spring gradually brings the bob to rest at A so that at the extreme end of the swing the speed for an instant is zero while it reverses. At this instant the spring is most bent and is therefore offering greater opposition than at any other instant during the motion.

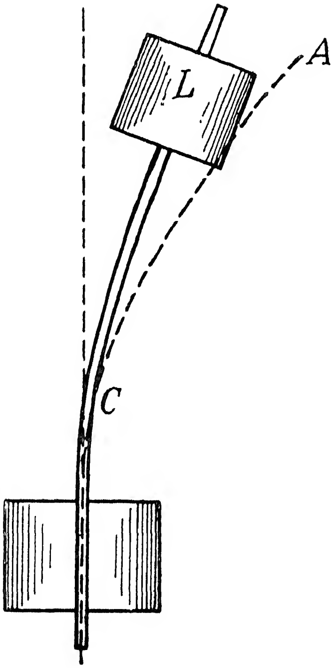


FIG. 123.

All this is represented to the eye for the quarter period we have been studying by the speed curve and the force curve shown in Fig. 124. In this figure velocity is shown positive because the velocity of the bob is towards the right; the force is shown negative because the action of the spring on the bob is towards the left. It is easy to understand that the momentum is brought to zero more quickly the less its original amount and the stiffer the spring. Therefore we can say that the time from centre to extreme, namely, the quarter period, is smaller with smaller mass L and with smaller length of spring C.

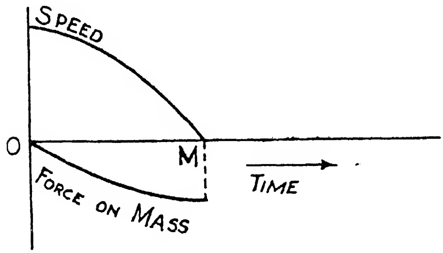


FIG. 124.

In the next quarter period the spring gives momentum to the mass which starts with very small velocity from A where the deflection is greatest. The speed increases so long as the spring pushes, that is to say until the centre is reached, when the spring's elastic action vanishes. We now get the portion MO′

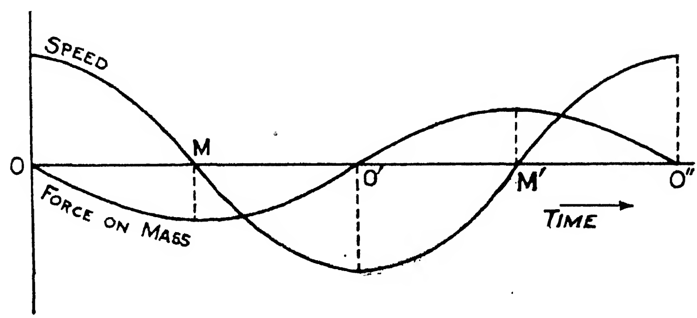


FIG. 125.

of the curves in Fig. 125, in which the velocity is shown negative because it is towards the left like the force.

The third quarter period begins with the bob at maximum left-ward speed and the spring straight and unresisting. As the bob moves, the spring is bent and begins to exert a force to the right. This continues with all the circumstances of the first

quarter period reversed until the maximum distance from the centre is reached. All the left-ward motion of the bob has been destroyed by the elastic force of the spring. The fourth quarter period is a repetition of the second quarter period reversed. The curves for a complete period are therefore as shown in Fig. 125.

Energy Relations.

98. It may be noticed at once that the energy of motion possessed by the bob at the middle of its swing is consumed in bending the spring, which thus stores it as potential energy. This potential energy is in its turn consumed during the next quarter period in creating kinetic energy, which is merely another kind of storage. And so the exchange recurs continually without any loss and in such a manner that at intermediate stages the sum of the amounts of energy in the two forms remains constant in accordance with the principle of the conservation of energy.

Let us decide upon such a form of spring that the sideways deflection of a strip C cm long is at the rate of C cm per dyne of sideways force whatever the value of C. Then when its deflection on any occasion is q cm we know that its elastic tendency towards its mean position is q/C dynes. In forcing it outwards from the mean position to a distance q the force required increases uniformly from 0 to q/C, its average value is $\frac{1}{2}q/C$, and therefore the work done is $\frac{1}{2}q/C \times q$ or $\frac{1}{2}q^2/C$. In other words, the potential energy of the spring when the bob is displaced a distance q is $\frac{1}{2}q^2/C$ ergs. The kinetic energy of a mass L grams moving with velocity i cm per second is $\frac{1}{2}Li^2$ ergs. Thus the statement that the sum of these energies is the same at any time during the motion may be written

$$\tfrac{1}{2}Li^2 + \tfrac{1}{2}q^2/C = \text{constant}.$$

At the extreme displacement Q the kinetic energy is zero and the potential energy reaches its maximum value $\frac{1}{2}Q^2/C$. While passing through the mean position the potential energy is zero and the kinetic energy has its maximum $\frac{1}{2}LI^2$. The constancy of the total energy gives the energy equations

$$\tfrac{1}{2}Li^2 + \tfrac{1}{2}q^2/C = \tfrac{1}{2}LI^2 = \tfrac{1}{2}Q^2/C.$$

99. Now the fact that the velocity and the force are in quadrature as regards their maximum and their zero values suggests

that (as an approximation at any rate) we might try if

$$i = I\cos\omega t$$

and

$$q = Q\sin\omega t,$$

which represents two sines in quadrature, with maximum values I and Q, can satisfy the energy equation. Trying these values we get

$$\tfrac{1}{2}LI^2\cos^2\omega t + \tfrac{1}{2}Q^2\sin^2\omega t/C = \tfrac{1}{2}LI^2 = \tfrac{1}{2}Q^2/C$$

or

$$\tfrac{1}{2}LI^2(\cos^2\omega t + \sin^2\omega t) = \tfrac{1}{2}LI^2$$

or

$$\cos^2\omega t + \sin^2\omega t = 1,$$

which is perfectly true for all values of t. Therefore the trial solution is a valid one. It can be shown that there is no other essentially different solution. The curves of Fig. 125 ought therefore to be drawn as sine curves.

Period.

100. We have already learnt that the average value of a quarter or half of one sine wave is $2/\pi$ times the maximum ordinate, and therefore we know that the average value of the velocity i is $2I/\pi$. The whole displacement Q is therefore covered in the time $Q \div$ average velocity, that is $\pi Q/2I$. By aid of the equation

$$\tfrac{1}{2}LI^2 = \tfrac{1}{2}Q^2/C$$

we obtain

$$\frac{\pi Q}{2I} = \frac{\pi}{2}\sqrt{(LC)}.$$

We have already called this interval of time $\frac{1}{4}T$. Therefore

$$T = 2\pi\sqrt{(LC)}$$

From § 1, III., we know that $\omega T = 2\pi$ and we deduce

$$\omega = 1 \div \sqrt{(LC)}.$$

The amplitude Q, it should be noticed, has disappeared from the formulæ, which shows that the period and the frequency are independent of the amplitude. This is a rather negative kind of proof of this important fact, so we may profitably examine the matter in another way.

Isochronism.

101. Consider two experiments with the same spring pendulum, the amplitude in one experiment being double that in the other. Fig. 126 represents them simultaneously at the extreme positions.

The restoring force in the case of A_2 is twice that of A_1, the velocity communicated to A_2 in, say, 1/1,000 of a second, is twice that given to A_1, the small distance covered by A_2 is therefore twice that covered by A_1, and consequently the distance of A_2 from its mean position remains twice the distance of A_1 from its mean position. This argument can clearly be applied in all successive intervals of time throughout the motion, which proves that A_2 always remains twice as far as A_1 from the mean position. Consequently the whole journey to the mean position is accomplished in the same time by both vibrators. The same argument

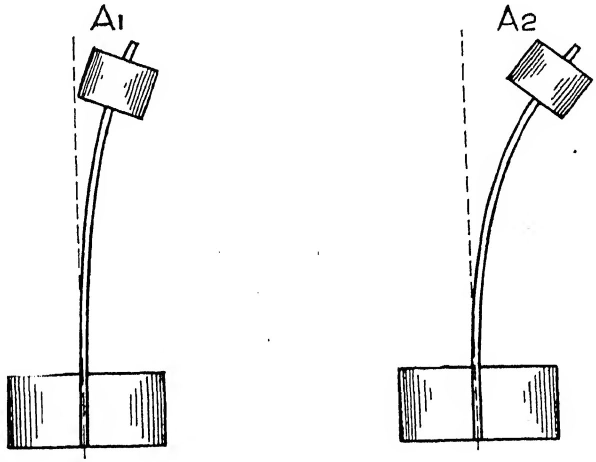


FIG. 126.

applies whatever multiple the displacement of A_2 may be of A_1. The quarter period is thus the same for all amplitudes. This is the principle of isochronism (equal times) first clearly stated by Galileo, and is said to be the fruits of his observations in church on the swinging of the censer.

Acceleration and Displacement.

102. A further fact emerges on closer examination of the motion of either vibrator. The mass of the bob is L grams, and therefore when its motion is seen to be accelerating we know that the force acting upon it must be L times the rate of increase of its velocity, which is L times the rate of increase of the rate of decrease of its distance from its mean position. But the force

acting on the bob comes from the spring, which exerts force according to the law

$$\text{force} = \text{displacement} \div C.$$

Therefore we conclude that

displacement $= LC$ times the rate of increase of the rate of decrease of itself.

Sine curves have the property expressed by this equation, and it can be shown that only sine curves have it. The matter is most easily discussed by putting q for displacement and d/dt for rate of increase. Then the above statement becomes

$$q = LC\frac{d}{dt}\left(-\frac{d}{dt}\right)q.$$

The general solution of this is

$$q = a\sin(\omega t + \phi)$$

because

$$\frac{dq}{dt} = a\omega\cos(\omega t + \phi)$$

and

$$\frac{d^2q}{dt} = a\omega^2\sin(\omega t + \phi)$$

whence

$$\frac{d^2q}{dt^2} = -\omega^2 q.$$

This general solution is evidently a solution of our special equation provided

$$\omega^2 = 1/LC.$$

This result agrees with that of § 100.

Since $\frac{d^2q}{dt^2}$ is the acceleration outward and $-\frac{d^2q}{dt^2}$ is the restoring acceleration we may write the differential equation in words thus :—

the square of the angular velocity = the restoring acceleration ÷ the displacement ;

or again :—

period $= 2\pi\sqrt{(\text{displacement} \div \text{restoring acceleration})}$,

in sine motion.

From the discussion we may also derive this important principle :—When a mass moves under a restoring force proportional to the displacement, the motion is a sine motion, that is, is simple harmonic.

Force and Speed.

103. In all that has preceded, the motion of and the force exerted upon the mass have been discussed and shown in the diagrams. If now the spring's point of view is in turn taken and its curves of force and speed drawn we get Fig. 127. In this the

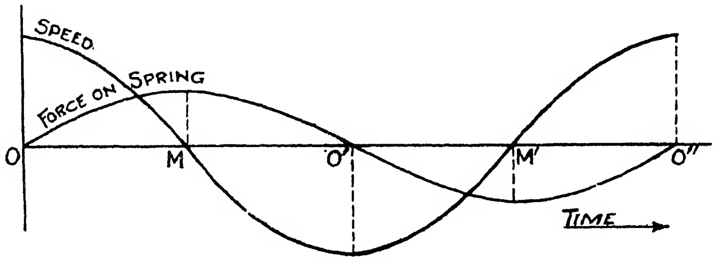


FIG. 127.

speed curve leads the force curve, in contrast to Fig. 125. This is because the force exerted on the spring by the inertia of the moving mass is equal and opposite to the force exerted on the mass by the spring—and it is the latter is shown in Fig. 125. These oppositely directed forces are merely the two aspects of the phenomenon. The speed curve is the same for both the spring and the mass.

Putting the diagrams into words we may say that the force on the spring lags a quarter period behind the speed of the spring;

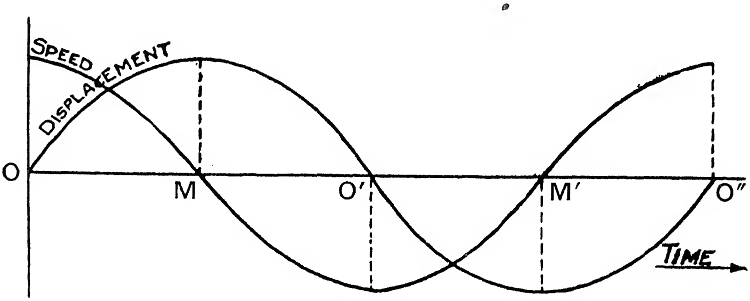


FIG. 128.

while the opposite is true for the mass, for the force on the mass leads the speed by a quarter period. The displacement of the spring and the mass are necessarily the same, like the speed. The curves of speed and displacement are drawn in Fig. 128.

Interpretation of the Analogy.

104. In slow mechanical vibrations the displacement is the thing most easily observed ; we watch the progress of the vibrations, if slow enough for the eye to follow, by the to-and-fro motion of the mass. In electrical affairs the things most easily chronicled are the current and the voltage, by means of ammeters and voltmeters. Current is analogous to the speed and voltage to the force in the mechanical case.

Displacement in the latter case is the analogue of electrical displacement or, what comes to the same thing, the quantity of electricity moved. Thus all the arguments set forth above are easily translated into electrical terms. If the reader will glance through them again making the substitutions indicated and replacing the spring C by a condenser of capacitance C, and the mass L by a coil of inductance L, it will be seen that the currents and voltages arising during the natural oscillation of a resistanceless closed electrical oscillator are explained and their relationships set forth quantitatively by the mechanical analogy.

Vibration with Friction

105. A spring and mass if once set going will vibrate a very long time if the frictional resistance is small, but as air friction can never be eliminated entirely, the amplitude will gradually diminish and the motion decay unless help be rendered in such a way as to compensate for the resistance. When the help afforded is in the nature of an applied force and is exactly equal at each instant to the resisting force the motion will, obviously, be absolutely unaffected—that is to say we shall have continuous sine oscillations. It is customary and nearly correct to assume, in mechanical problems, as a first approximation, that the frictional resistance is proportional at each instant to the speed, say, R times the speed. The multiplier R is called the " coefficient of resistance." On the electrical side of the analogy the term is shortened to " resistance " and the assumed relationship between friction and speed becomes the very accurate statement known as Ohm's Law. Applying now the reasoning given above, we see that if the vibration is to be kept of sine type in spite of friction a force R times the speed must be applied at each instant to the

vibrator. Since the friction is in the opposite direction to the movement, the compensating force will be in the same direction as the movement. Therefore in drawing the curves of Fig. 129 this applied force is represented by a curve whose ordinates are R times those of the speed curve. It is a sine curve having zeroes and maxima at the same times as the speed curve ; in other words, the external force needed to maintain sine vibrations is a sine force of frequency equal to that of the *frictionless* vibrator, of amplitude sufficient to compensate for the friction, and in phase with the speed.

Perhaps the most important of the above deductions is the contrast between this force needed to overcome the resistance and that exerted by the spring to keep the mass moving in the frictionless case. The latter leads the speed by a quarter period ; the

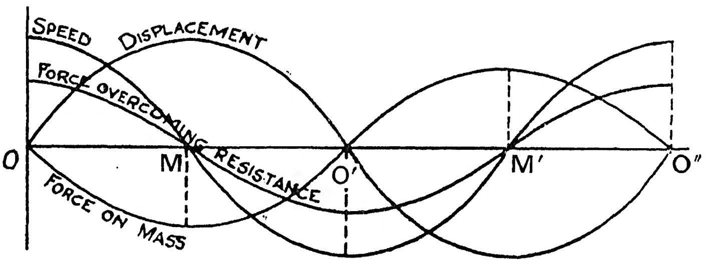


FIG. 129.

former is exactly in phase with the speed, just as if the frictional resisting force existed alone. In fact, looking at the phenomenon as an alternating applied force working against an alternating resisting force we see that the spring and the mass might as well be non-existent. Towards the applied alternating force occupied with the resistance—alternating at the frequency natural to the frictionless vibrator—the vibrator behaves as if it had neither inertia nor elasticity ; in other words, the inertia and the elasticity cancel each other at the particular frequency indicated.

106. It is important to emphasise that the frequency spoken of in the above paragraph is not the natural frequency of the vibrator moving under friction. Of course, in the case of a vibrator like that in Fig. 122, the air friction may be so feeble that there is practically no difference between the actual frequency and the ideal frictionless frequency ; but if the air friction were increased

by fitting paper damping vanes on the bob, a noticeable change of the free frequency would arise. By the free or natural frequency is meant that actually observed in the absence of a sustaining external force—it is the frequency of the dying oscillation. From § 47 it can be shown that the natural frequency of a vibrator is reduced by the presence of resistance by an amount proportional to the square of the resistance, provided damping is small.

107. It is possible to apply to a vibrator an alternating force of any frequency whatever, but there are clearly two cases of special interest. One is the case already discussed in which the applied force is of the frequency that the vibrator would have in the absence of friction. The force and the vibrator are then said to be "in resonance." The other case is that in which the applied force is of the frequency natural to the vibrator. The force is then said to be "in synchronism" with the vibrator. Unless the damping is very great, resonance and synchronism are almost the same and both cases are regarded as included in the vague expression "the vibrator is tuned to the force."

Forced Vibrations.

108. When an external sine force acts upon a vibrator, whether at the resonance frequency or not, it gradually builds up a vibrating motion to a definite final amplitude and sustains it thereafter at that amplitude. The frequency of this final sine motion is always the same as that of the force, and the motion is called a forced vibration. We have seen that in the case of resonance the elasticity and the inertia cancel each other and the force (in this case only) has no other occupation than overcoming the resistance. Since the resisting force is Ri, in the notation already adopted, if E be the applied force we have $E = Ri$, and the force, it has already been seen, is in phase with the velocity. The final velocity, therefore, is $i = E/R$ in the case of resonance. We are immediately confronted with the question: What happens when, the force being kept of constant frequency, the vibrator has its inertia decreased to a value smaller than before? One thing is at once evident, namely, that all the springiness cannot now be cancelled since the inertia will not be large enough.

Massless Case.

109. Perhaps at this stage, while our endeavour is to obtain an understanding of vibration and not a collection of mathematical results, the question is best answered by going to extremes. Suppose the mass to be removed altogether, the spring to possess negligible inertia of its own, and friction to be absent. The sine force is applied to the end of the spring, and the spring responds to the force instantly because of the absence of inertia. Starting from the mean position let us trace the circumstances of an outward journey. The spring merely behaves as a record of the growth of the force, for its deflection is at each instant proportional to the magnitude of the force—it is at the mean position when the force is zero, at its extreme displacement when the force is at

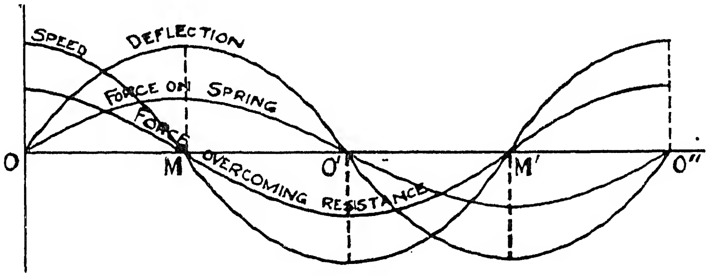


FIG. 130.

the maximum value. Therefore the deflection of the spring can be represented as a sine curve with its ordinates proportional to those of the force curve and in phase with this curve. If we take Q to be as before the maximum amplitude of the spring's motion, the maximum value of the force is Q/C, where C is the length of the standard spring. The rate of increase of the deflection curve gives the speed curve, which is therefore a sine curve leading the displacement curve by a quarter period. The curves are shown in Fig. 130.

In this massless and frictionless motion the function of the spring is to take up or occupy the driving force; and the work done by the latter on the outward journey is stored by the spring's resilience and yielded up again on the inward journey. In consequence the driving force does no work on the whole and the motion may be called "wattless" or "idle." This wattless force is in quadrature with the speed and lagging.

110. Now suppose that resistance is introduced and that this force of resistance is proportional to the speed, say Ri. By multiplication of the speed ordinates by R we get the fourth curve of Fig. 130. This last curve is drawn of such sign as to represent the force required to overcome the resistance. Thus to maintain the motion a new force is needed which is in phase with the speed and therefore a quarter period ahead of the wattless force. Its amplitude is RI, if I is the maximum speed in the forced vibration. As the force of resistance opposes the motion on both inward and outward journeys and on each side of the mean position, work is required at every point of the journey to overcome it. The new component added to the applied force may be called the "power" or the "energy" component, or indeed the "load" component, if we regard the friction as the load upon the system.

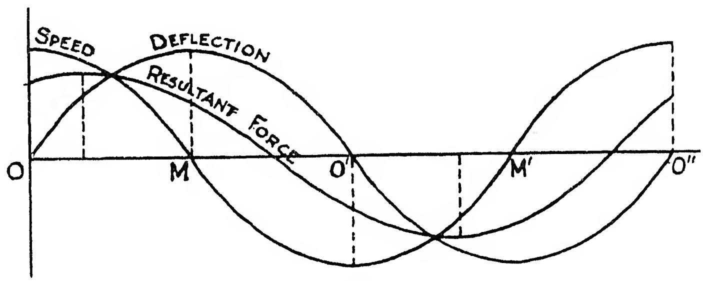


FIG. 131.

The sum of two sine curves is another sine curve in an intermediate position (§ 5), and thus the two components of the applied force may be combined into a single sine curve of force. This is shown in Fig. 131. The single resultant is seen to be lagging behind the speed curve. The actual amount of the lag depends on the amplitudes of the component forces. The common intersection of the three curves is accidental.

Springless Case.

111. Take now the other extreme when, instead of removing the mass and leaving the spring, the spring is removed. Our experiments are supposed to be performed where there is no gravity, so that the mass must be supposed to be a lump of inertia without material support; or if that is too great a strain on the imagination, the mass may be supposed to be placed on a perfectly

smooth horizontal table. Let us suppose provisionally that the motion ultimately established when the alternating force acts upon the mass is a sine motion of which the speed curve is represented in Fig. 132. Starting from the mean position with

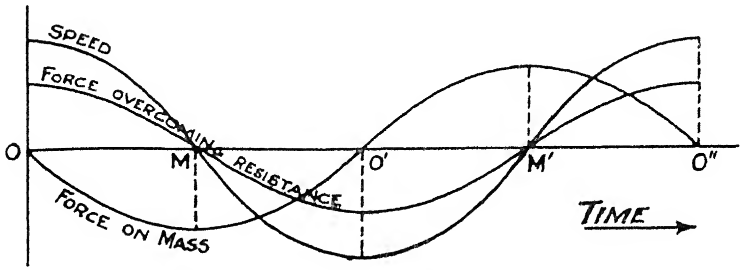


FIG. 132.

maximum speed we see generally that the speed is being reduced during all the outward journey and that therefore the force must be opposing the motion. The force is in fact destroying the momentum and is equal at any instant to the product of mass and acceleration. At the beginning of the quarter period the speed curve is level and is undergoing no change, the acceleration

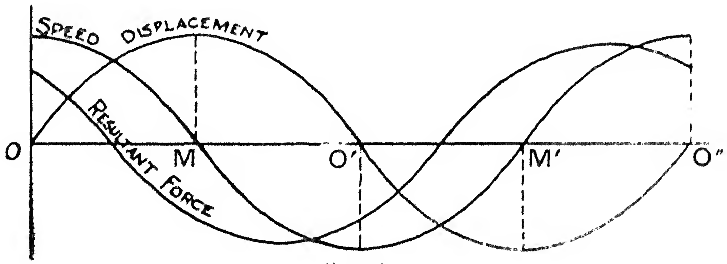


FIG. 133.

is zero and therefore the force is zero. At the end of the stroke the speed curve is steepest, the acceleration or rate of change of speed is greatest, and there the force is greatest and has the value $L \times$ maximum acceleration. From the definitions of acceleration, speed, and displacement it is now clear that if the force is given as a sine force, the acceleration, the speed and also the displacement curves must be sine curves. Force and speed appear in Fig. 132. They agree with Fig. 129, where the force applied to the mass happened to come from a spring instead of from

the external source assumed here. If the speed is, as before, $i = I \cos \omega t$, the maximum value of the force is $L\omega I$ since the maximum acceleration is equal to ωI.

The function of the mass is to occupy the force, which during the inward journey gives kinetic energy to the mass and during an outward journey takes it back. On the whole no work is done. The force is in leading quadrature to the speed and is "wattless."

112. If the motion of the mass is now opposed by friction proportional to the speed at any instant, an additional sine force must be introduced to overcome the friction in order to leave the other component force and its duties unaltered. This new component force has a curve of which all the ordinates are R times the speed ordinates. It is therefore in phase with the speed curve and in quadrature with the wattless component force. Its maximum value is RI and it is called the load component. When combined with the wattless component as in Fig. 133, the resultant leads the speed curve by less than a quarter period.

113. The results now reached in Figs. 130, 131, 132, 133, may be put in the following broad way: When a sine force is put to work moving an object whose motion is resisted by friction proportional to the speed, the steady motion ultimately attained has a sine for its speed curve, which leads the force curve if the object exhibits only elasticity, and lags behind if the object exhibits only inertia. This phase difference is due in each case to a quadrature component taken from the applied force to overcome the springiness or the massiveness, as the case may be, which is led by the speed in the former case and leads the speed in the latter. If the speed equation is

$$i = I \cos \omega t$$

the maximum displacement is $Q = I/\omega$, and the maximum acceleration is ωI. Therefore the quadrature wattless components have respectively the maximum values

$$I/C\omega \text{ and } L\omega I,$$

if C is the length of the standard spring, L the inertia of the mass. In either case the remaining component of the applied force has the maximum value RI, is in phase with the speed, and works against the friction at the average rate $\frac{1}{2}RI^2$.

General Case.

114. We may now return from these extreme cases to the problem of the mass and spring acted upon by an alternating force of the old frequency, the mass being too small to bring the vibrator into resonance. Under such a condition not all the springiness is cancelled (though all the inertia is), and therefore curves of the kinds given in Figs. 130 and 131 are to be expected—the speed will lead the force. In order to calculate the amplitude of all the curves we shall have to know how to estimate the uncancelled residuum of springiness. This is suggested by the results of § 113. For the amplitude of the wattless force curve is given by

$$\frac{I}{C\omega} - L\omega I$$

since the spring component leads the inertia component by two right angles. The measure of the uncancelled springiness is therefore

$$\frac{1}{C\omega} - L\omega.$$

This is not zero, it should be noticed, because we have supposed

$$\omega^2 \text{ not equal to } 1/LC,$$

that is to say L and C have not values that give resonance.

115. Suppose now finally that with the old value of the mass the spring were lengthened so as to become weaker. Clearly the inertia cannot now all be cancelled. We should therefore expect such motion as is indicated by Figs. 132 and 133. The speed of the steady motion would lag behind the applied force. The uncancelled inertia would have the effect

$$L\omega - \frac{1}{C\omega}$$

and the component force devoted to keeping it moving would be

$$\left(L\omega - \frac{1}{C\omega}\right) I.$$

It should be noticed that in passing gradually from a state where the springiness is in excess to one in which the massiveness is in excess, we must take the vibrator through the resonance adjustment, in which $L\omega = 1/C\omega$. During that transition the speed curve moves from a leading phase to a lagging one, passing through the cophasal state at resonance.

In the cases of §§ 114, 115 the applied force must also supply a component having the amplitude RI for overcoming the friction.

Mnemonical Experiment for the Reader.

116. Many of the above facts can be recalled to mind by performing a series of simple experiments needing as apparatus merely a piece of thread and a weight. Tie one end of the thread to the weight and the other to a fixed support and set the bob swinging. The motion is practically a frictionless sine motion

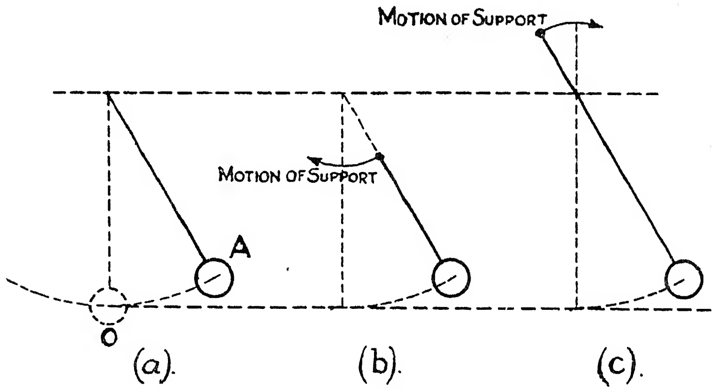


FIG. 134.

which is brought about because gravity tends to bring the bob towards O, Fig. 134 (a), pulling the bob more strongly towards O the further it is displaced from O. At O the displacement of the bob is zero, its speed is maximum, and the pull of gravity along the path of the bob is zero. At A the displacement is maximum, the speed is zero and the applied force is a negative maximum. Similar results can be obtained at other important points of the swing and the deduction drawn that in free vibration the active component of gravity is in opposite phase to the displacement and leads the speed by a quarter period.

Take the upper end of the thread in the hand and move the hand to and fro in the place of the motion and at the same frequency. A large amplitude can be worked up because of the resonance. Now shorten the thread a little and move the hand at the same frequency as before. A forced vibration of the

frequency of the hand will be started and the pendulum in fitting itself to the circumstances will move as if part of a longer pendulum, Fig. 134 (b). Now lengthen the string and again oscillate the hand at the same frequency as before. The pendulum will oscillate as if shorter than its real length as indicated in Fig. 134 (c). The relative phases of the motion of the hand and of the bob and of the gravity component force are left to be determined by the experimenter.

CHAPTER IV

COUPLED CIRCUITS AND TRANSFORMERS

1. When two circuits are so arranged that current or emf in either produces an immediate current or emf in the other they are

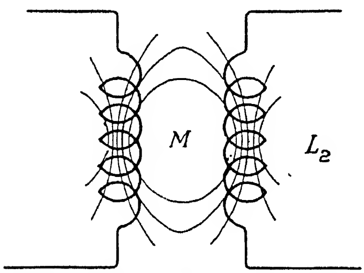


FIG. 135.—Indirect Magnetic Coupling.

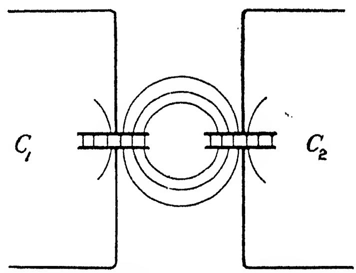


FIG. 136.—Indirect Electric Coupling.

said to be coupled. The thing doing the coupling may be a magnetic flux, an electric strain, or the difference of potential that arises when current traverses a resistance. Thus we have magnetic,

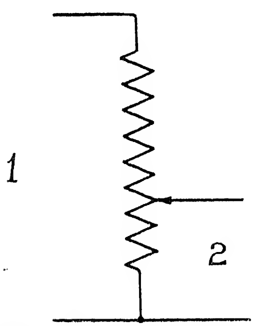


FIG. 137.—Ohmic Coupling.

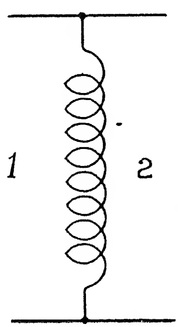


FIG. 138.—Direct Magnetic Coupling.

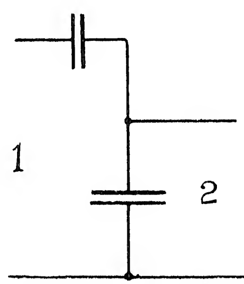


FIG. 139.—Direct Electric Coupling.

electric and ohmic coupling. When the coupling is effected without material conducting connection it is said to be indirect, but when effected otherwise it is called direct. Fig. 135 shows indirect magnetic coupling, the flux produced by the coil marked

L_1 threading, in part, that marked L_2. In Fig. 136 we have indirect electric coupling ; some of the Faraday lines due to presence of electricity in one circuit spread from the condenser marked C_1 to the condenser C_2. A number of illustrations of various kinds of direct and indirect coupling are added in Figs. 137 to 142.

2. Power is usually supplied to one circuit and consumed in or disposed of by the other. The former is called the primary

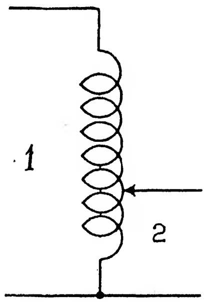


FIG. 140.—Direct Magnetic Coupling.

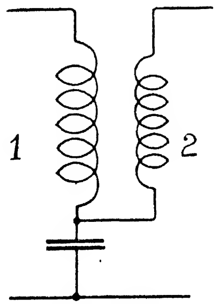


FIG. 141.- Mixed Magnetic and Electric.

and the latter the secondary circuit. Moreover, as these paired circuits are utilised for modifying the power supplied as regards its voltage or current magnitude so as to be more suitable for the consumer, the coupling portions are commonly called transformers. Resistance coupling, which is in its essence direct, is distinguished from the other kinds by the fact that power can be transferred whether the current is continuous or varying. Including this, we may say that power can be continuously transferred in a regular manner in all cases with alternating current. In what follows we shall assume that the source of power in the primary circuit is an alternating emf of sine form.

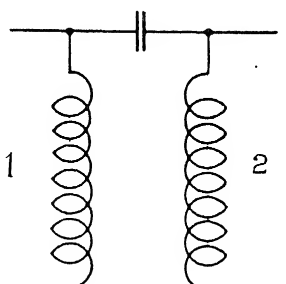


FIG. 142.—Mixed Magnetic and Electric.

DIRECT COUPLING

It is convenient first to discuss direct coupling, for its theory is only an extension of that of two conductors placed in parallel, which has been fairly fully treated in the last Chapter.

Resistance Coupling with Ohmic Load.

3. A typical case of resistance coupling is shown in Fig. 143. The assemblage may be regarded as comprising three distinct circuits, namely, $ORAR_{\prime\prime}E$ and $OR_{\prime}AR_{\prime\prime}E$ and $ORAR_{\prime}O$. Of these circuits two contain the source E and one does not. We therefore call the last one the secondary circuit; and we imagine the primary to include either R or $R_{\prime}$. There are usually practical circumstances or historical associations that decide on one of these being taken as the primary circuit. For instance R may be part of the same winding as $R_{\prime\prime}$, while $R_{\prime}$ is a distinct resistance variable at will by the experimenter. We shall take $ORAR_{\prime\prime}E$ to be the primary circuit, so that R becomes the common part of the primary and secondary circuits.

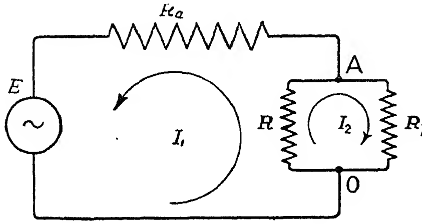


FIG. 143.

Notation.

4. For the purpose in hand the cyclic notation of the last Chapter is very convenient. As before, the sign of the cyclic currents is defined so that they are of the same sign when they are at any instant passing in the same direction through the common part R. This is indicated in the Figure by the bent arrows; the directions shown will be taken as positive. At an instant when a current is flowing in the other direction, it is called negative. These arrows must not be interpreted as meaning that the currents are ever simultaneously flowing in the directions shown nor that they are always flowing in those directions.

We shall indicate the total resistance in the primary circuit by the symbol R_1 and that in the secondary circuit by R_2. The resistance common to both circuits is called R. To help to decide the amount to be included in R_1 imagine the secondary circuit to be cut anywhere except in the part common to both circuits; then the whole resistance passed through in circulating round the primary in Fig. 143 is $R_1 = R + R_{\prime\prime}$. Similarly by cutting the primary it is easy to see that $R_2 = R + R_{\prime}$.

It is clear that the current leaving the alternator E is i_1 and that this passes through R_a and combines vectorially with i_2 when passing through R. It is clear too that the resistance R_b is traversed only by the current i_2. Our positive direction is, as it happens, opposite from what would really be produced by positive primary current, and therefore i_2 will usually be accompanied by a minus.

Primary and Secondary Currents.

5. For calculating the currents and the various properties of the circuits we may appeal to the shunt rule of § 67. Chap. III. Since $R_2 = R_b + R$ we obtain for the parallel resistance the value $R(R_2 - R)/R_2$. To obtain the whole resistance offered to e we must add to this expression the resistance R_a, which is equal to $R_1 - R$. On simplifying we obtain for the equivalent resistance the expression $R_1 - R^2/R_2$. Now the current flowing from the source is determined by the magnitude of the emf and by the equivalent resistance offered to that emf. Hence

$$i_1 = \frac{e}{R_1 - R^2/R_2}.$$

The current i_2 can now be calculated by aid of the results of III., § 67 for parallel circuits. In the present application it must be noted again that the direction of i_2 marked as positive in Fig. 143 is opposite to the actual direction of flow of current in the secondary circuit, and therefore we write in the formula just quoted $-i_2$, not $+i_2$. Hence

$$\frac{-i_2}{i_1} = \frac{R}{R + R_b} = \frac{R}{R_2},$$

and therefore

$$i_2 = -\frac{R}{R_2} i_1.$$

The minus arises merely because we defined the positive direction of i_2 as we did. If it be required, the current in the common branch R can easily be obtained either by the same rule or by noting that it is equal to $i_1 + i_2$ which equals $i_1 - (R/R_2)i_1$.

Leakage and Coupling Coefficients.

6. The effective resistance of the assemblage to its cyclic current i_1 is $R_1 - R^2/R_2$ which equals $R_1(1 - R^2/R_1R_2)$. Thus, the whole resistance of the primary is reduced in the proportion $1 : 1 - R^2/R_1R_2$ by the presence of the secondary.

The quantity $(R^2/R_1R_2)^{\frac{1}{2}}$ is called the coefficient of coupling. The quantity $1 - R^2/R_1R_2$ is called the leakage coefficient. The leakage coefficient has the symbol σ and the coupling coefficient the symbol k. Evidently $\sigma = 1 - k^2$. It is clear that k is always less than 1 because R_1 is greater than R which is a part of it, and R_2 also is greater than R for the same reason. Hence σ is always positive and less than 1.

Resistance Coupling with Reactive Load.

7. The circuits to be discussed under this heading are such as frequently arise in the application of the potential divider and the potentiometer to reactive apparatus.

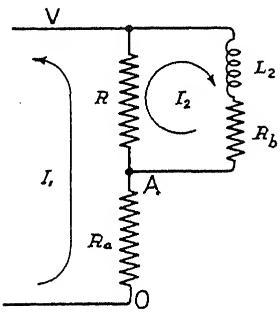


FIG. 144.

Taking the currents to have their positive directions as marked in Fig. 144, we may reason as follows in constructing the vector diagrams for the primary and secondary circuits. In the secondary the current i_2 produces potential drops of amplitude RI_2 and R_bI_2 in the resistances, and a potential drop $L_2\omega I_2$ in the inductance. These are perpendicular and the latter leads the former. We therefore obtain the vectors BD, DA in Fig. 145. The total potential drop in the secondary circuit is made up for by the voltage supplied to the circuit by the primary current i_1 traversing R. This emf is of amplitude RI_1 and tends to send current round L_2 and R_b in opposite sense to arrow I_2. We therefore show in the Figure the vector BA equalling the sum of the two sides BD, DA of the triangle. In the primary the current i_1 produces potential drops R_aI_1 and RI_1 in amplitude, set off to scale at OA, AB. The total drop is made up by the vector sum of the applied emf of amplitude E and the reaction from the

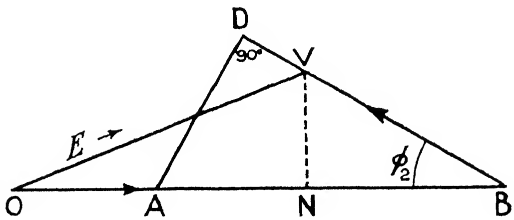


FIG. 145.

secondary due to i_2 flowing in R, which is of amplitude RI_2 in direction VB. The only way of fitting the voltages in is to take vector OV to represent the applied emf and to read the drawing as stating that applied emf OV and secondary reaction VB make up the primary drops OA and AB. Thus we conclude that an emf of phase and magnitude represented by crank OV is needed to send current of amplitude I_1 through the assemblage.

Ratio of Currents.

8. The ratio of the currents may be obtained directly from electrical considerations. For the voltage between A and V due to the current of constant amplitude I_1 is RI_1 and this tends to send current through R_b and L_2 in opposite direction to that taken as positive for the secondary current. We therefore write

$$I_2 = -RI_1/(Z_2, \phi_2)$$

where Z_2 is the impedance and ϕ_2 the phase angle of the whole of the secondary circuit, that is $\phi_2 = \angle L_2\omega/R_2$.

This equation may also be obtained from the crank diagram by expressing AB in the two possible ways. It means that numerically

$$I_2/I_1 = R/Z_2$$

and that the secondary current in the marked direction leads the primary by the angle $\pi - \phi_2$.

Equivalent Values.

9. From Fig. 145 it is seen that the equivalent resistance is ON and the equivalent reactance NV. The formula is given by

$$R'I = \text{OA} + \text{AB} - \text{NB}$$

$$\therefore R' = R_a + R - R\cos\phi_2\,(I_2/I_1)$$

$$= R_1 - R_2\,(R/Z_2)^2$$

Also

$$X'I_1 = \text{NV} = RI_2\sin\phi_2$$

$$= RI_2\,(L_2\omega/Z_2)$$

$$X' = L_2\omega\,(R/Z_2)^2.$$

Variable Reactive Load.

10. If we imagine the reactance $L_2\omega$ to vary while all the resistances and the primary current remain constant, the secondary current and the phase angle ϕ_2 must vary. At every

adjustment, however, the point D in Fig. 145 will lie on a semicircle of diameter AB. Since BV/BD, being equal to R/R_2, is

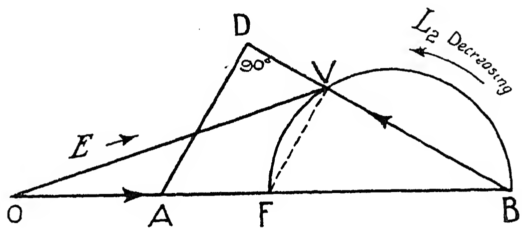


FIG. 146.

constant, the point V will describe a semicircle of diameter BF (Fig. 146), where BF : BA $= R : R_2$. The direction of decrease

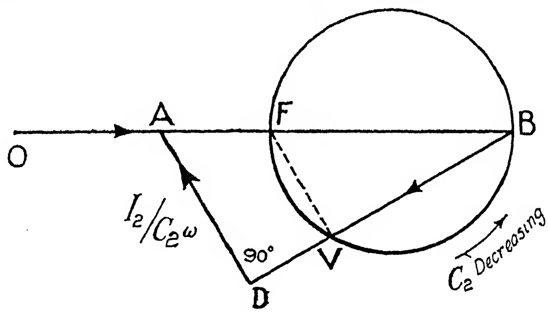


FIG. 147.

of L_2 is marked. The height of V above OB indicates the value of $X'I_1$ which is proportional to X' since I_1 is constant, and this is seen to have a maximum value when $\phi_2 = 45$ degrees, that is when $L_2\omega = R_2$.

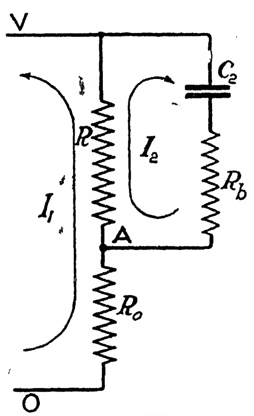


FIG. 148.

11. If the point V be supposed to pass across the line OB from the upper semicircle to the lower, Fig. 146, the sign of $L_2\omega$ changes, that is to say the reactance becomes negative. In other words, the lower semicircle relates to circuits wherein a condenser replaces the inductance of the preceding problem. This is another illustration of the fact that in any formula or chain of reasoning $L_2\omega$ may be replaced by $-1/C_2\omega$ with confidence in the

results. The formulæ obtained above may be modified by the reader so as to give the solution of the problem of Fig. 148.

General Remarks on Transformers

The instances of resistance coupling just discussed are not usually included in the transformer class but are called potential dividers or potentiometers. The term transformer is usually reserved for the type of apparatus in which the two circuits are linked by a common magnetic or electric flux. This mutual flux, which is analogous to a gearing, acts as a vehicle of energy and performs its duties with very small loss.

Reactive Transformers.

12. In the discussion of reactive transformers there are three very useful quantities frequently arising, namely, the ratio of the secondary current to the primary, the equivalent resistance, and the equivalent reactance. The latter have already been defined in dealing with shunt circuits (III., § 83) and given the symbols R' and X'; from them can be obtained the impedance Z' and the phase angle ϕ' of the assemblage. We shall represent the total primary reactance by X_1 of which there may be a part X common to both circuits and a part X_a not common. Thus $X_1 = X_a + X$. Similarly the total secondary reactance will be represented by X_2 and the parts by X and X_b, so that $X_2 = X_b + X$ (see Fig. 149). Even if there be no common portion of the circuits and $X_a = X_1$ there may still be certain flux linkages, magnetic or electric, between the circuits, and we may represent the effect of the primary on the secondary by $X_m I_1$ and of the secondary on the primary by $X_m I_2$.

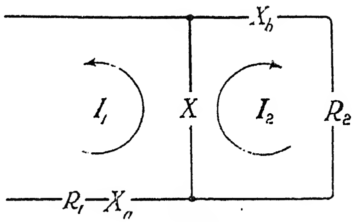


FIG. 149.

Ratio of Currents.

13. The portion of the reactance common to both circuits is traversed by the primary current I_1, and therefore its electromotive effect in the second is XI_1. Now the impedance of the

secondary may be represented as in the § 8 by (Z_2, ϕ_2) and then we may write

$$I_2 = XI_1 \div (Z_2, \phi_2).$$

Thus the current ratio, dropping the phase difference, is

$$I_2/I_1 = X/Z_2.$$

If the circuits are not linked by common reactance but merely by mutual flux linkages we need only put X_m instead of X in the above to represent the facts briefly.

Equivalent and Image Resistance.

14. The power delivered to the secondary circuit is consumed, usually, by the resistance R_2 present there. The rate of dissipation is $\frac{1}{2}R_2I_2{}^2$. When we replace the whole transformer by an equivalent single circuit carrying the primary current this rate of consumption must be mimicked by a portion of the single circuit—a portion over and above the primary resistance—which we shall call the image of the secondary circuit in the primary and represent by the symbol R_2'. The image traversed by the primary current I_1 must dissipate energy at the same rate as the actual resistance does when traversed by the secondary current.

Therefore $$\tfrac{1}{2}R_2'I_1{}^2 = \tfrac{1}{2}R_2I_2{}^2.$$

Hence the image resistance is

$$\begin{aligned} R_2' &= (I_2/I_1)^2R_2 \\ &= (X/Z_2)^2R_2, \end{aligned}$$

by aid of the last paragraph.

Having determined the image resistance of the secondary, the equivalent resistance of the whole transformer is

$$\begin{aligned} R' &= R_1 + R_2' \\ &= R_1 + (X/Z_2)^2R_2. \end{aligned}$$

This result is valid also when the common reactance X is replaced by a mutual reaction X_m established by flux linkages.

Equivalent and Image Reactance.

15. The inductance and capacitance of a circuit act in different ways as storehouses for energy, which flows in and out without dissipation in the ideal case. This function of the reactance in the secondary circuit must be imitated precisely by the image of the secondary reactance in the primary. We have already seen that if any current of form $I_2 \sin(\omega t + \phi_2)$ flows through a

reactance X_2 the P.D. at its terminals is $X_2 I_2 \cos(\omega t + \phi_2)$. Therefore the rate of working in the reactance is

$$X_2 I_2{}^2 \sin(\omega t + \phi_2) \cos(\omega t + \phi_2)$$

or

$$\tfrac{1}{2} X_2 I_2{}^2 \sin 2(\omega t + \phi_2).$$

This tidal wash of energy through the reactance, though it averages out at zero, has an amplitude $\frac{1}{2}X_2 I_2{}^2$ and this must be equated to $\frac{1}{2}X_2' I_1{}^2$, where X_2' is the image of X_2. Thus

$$X_2' = (I_2/I_1)^2 X_2$$
$$= (X/Z_2)^2 X_2.$$

It will be seen later that the effect of positive values of X_2 is to subtract from the primary reactance, while the effect of negative secondary reactance is to add to the primary reactance, in making up the equivalent reactance. We therefore write for the equivalent reactance

$$X' = X_1 - X_2' = X_1 - (X/Z_2)^2 X_2.$$

The above result is valid also when the common reactance X is replaced by a mutual reactance X_m.

16. It is a useful exercise on the operator method to obtain these results by generalisation of the simple formula obtained in § 5 for resistance coupling. It is only necessary to write $R_1 + X_1 D/\omega$ instead of R_1, XD/ω instead of R, and $R_2 + X_2 D/\omega$ instead of R_2 in the formula

$$R_1 - R^2/R_2.$$

On multiplying out after the substitutions indicated and arranging the terms we get

$$R_1 + (X/Z_2)^2 R_2 + \{X_1 - (X/Z_2)^2 X_2\} D/\omega,$$

which must be equated to

$$R' + X'D/\omega.$$

Many of the problems discussed below are particular cases of the general case of the preceding five paragraphs, but they will be treated separately and fully by the crank diagram method, in order to exhibit clearly phase relationships and other special features of the voltages and currents.

Direct Magnetic Coupling

17. A typical circuit is given by Fig. 150. We construct the diagram of Fig. 151 as follows: Draw OA in any convenient direction to represent the primary ohmic potential drop to scale. Set off AB a right angle in advance to represent $L_a \omega I$ to the same

scale, and BD in the same line to represent $L\omega I_1$. At B make an angle DBF equal to ϕ_2, the secondary circuit's phase angle. From D drop a perpendicular to determine F and join DF. Then DF represents $L_2\omega I_2$ and FB represents $R_2 I_2$ to scale. From D mark off DV the same proportion of DF as L is of L_2 and join OV. Then DV equals $L\omega I_2$ and VF equals $L_b\omega I_2$. The reaction of the secondary on the primary circuit is represented by the voltage DV, and this adds geometrically to the cranks OA and AD to give the applied emf in the primary, namely, OV $= E$. This is the emf required to force a constant current I_1—rms or amplitude—through the transformer primary with the given load. It is in fact the equivalent impedance of the whole transformer if we care to suppose $I_1 = 1$. It should be noted that the reaction DV from the secondary is in the present instance such as to tend to make the equivalent reactance of the transformer less than the primary reactance and the equivalent resistance greater than the primary resistance. This is more clearly seen by drawing VN and AN to exhibit NV the equivalent reactance and AN the image resistance. The image reactance is DN′, which is obtained by drawing VN′ perpendicular to AD.

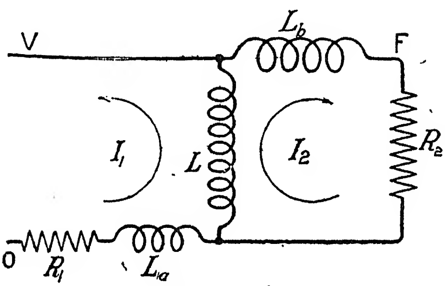


Fig. 150.

The Current Ratio.

18. The ratio of the primary to the secondary current is exhibited to the eye on this diagram by the cranks DV and BD, for the former is $L\omega I_2$ and the latter $L\omega I_1$. But as these cranks give inductive potential drops the currents are in fact a right angle behind them in phase. The phase relations, though not the relative magnitudes of the currents, are best seen either in the cranks FB and OA, the former being $R_2 I_2$ and the latter $R_1 I_1$ or in the cranks VK and OA of Fig. 152.

The formula of the current ratio may of course be deduced from the crank diagram. It is only necessary to express BD

in its two alternative ways; first as $L\omega I_1$ and then as $(R_2^2 + L_2^2\omega^2)^{\frac{1}{2}} I_2$ or $Z_2 I_2$. Thus we obtain again (see § 13)

$$I_2/I_1 = L\omega/Z_2$$

neglecting the phasal relations.

Variation of Load

Resistance Load Variation.

19. From V in Fig. 151 draw a parallel to FB determining the point K in Fig. 152. Then

$$\begin{aligned} \mathrm{DK/DB} &= \mathrm{DV/DF} \\ \therefore \mathrm{DK} &= L\omega I_1 \,.\, L\omega I_2/L_2\omega I_2 \\ &= (L^2/L_2)\omega I_1 \end{aligned}$$

Also

$$\begin{aligned} \mathrm{DK/DA} &= \mathrm{DK}/L_1\omega I_1 \\ &= L^2/L_1L_2. \end{aligned}$$

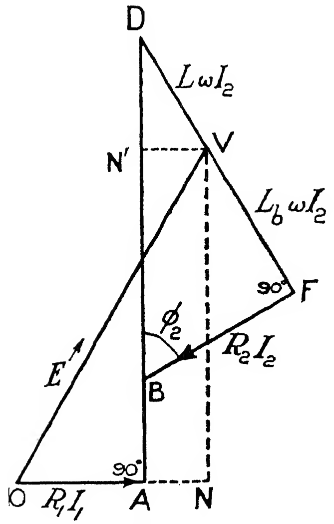


FIG. 151.

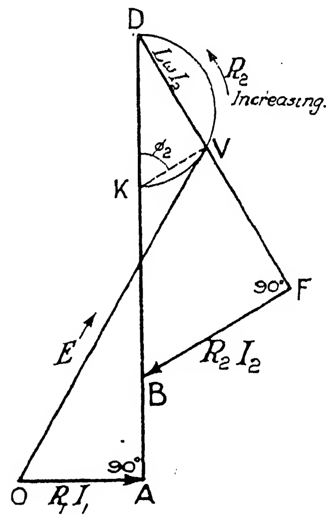


FIG. 152.

Now L/L_1 is the proportion of the coupling inductance to the whole of the primary inductance and L/L_2 is the analogous proportion for the secondary circuit. The geometric mean of these proportions is called the coefficient of coupling k. Thus the point K divides the crank AD in the proportion k^2 to $1-k^2$. The latter is called the leakage coefficient σ. We have therefore DK representing k^2 and AK representing σ in the crank diagram.

So long as the inductances are constant the point K is fixed however R_2 and R_1 vary, and if I_1 is constant KD is constant.

When the load R_2 varies the angle ϕ_2 varies; as R_2 increases the angle decreases, KV increases, and VD decreases. The angle KVD is always a right angle, and therefore V moves on a circular arc towards D as R_2 increases. The consequent changes in the effective impedance of the whole transformer are easily traced by watching the vector OV. When $R_2 = 0$, V is at K and OK is the primary voltage then. When $R_2 = \infty$, V is at D, the secondary is open and the voltage $E = $ OD is the same as if the primary were alone. The effective impedance is greatest when OV bisects KD.

Numerical Example, Fig. 153.

20. Let $R_1 = 2\,\Omega$, $L_a = 1$ H, $L = 4$ H, $L_b = 0{\cdot}02$ H, $f = 50 \sim$, $I_1 = 1$ A. Here L_a may be taken to include the inductance and R_1 the resistance of the alternator. Then $\omega = 314$, OA $= 2$ V, AB $= L_a\omega I_1 = 314$ V, BD $= L\omega I_1 = 4 \times 314 = 1{,}257$ V, $k^2 = L^2/L_1L_2 = 0{\cdot}795$, KD $= k^2$AD $= 0{\cdot}795 \times 1{,}571 = 1{,}246$.

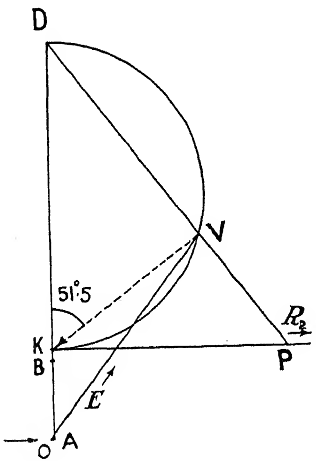


FIG. 153.

First take the secondary resistance $R_2 = 1{,}000\ \Omega$.
Then $\tan\phi_2 = L_2\omega/R_2 = 1{\cdot}26$.
Therefore $\phi_2 = 51{\cdot}5$ degrees.

Set off this angle at DKV, draw a semicircle on KD as diameter, draw KP perpendicular to KD, prolong DV to cut KP at P. Mark P 1,000 Ω. Then if KP be graduated uniformly so that KP $= 1{,}000$ units, the position of V is found for any value of R_2 by finding the corresponding number on these graduations and joining it to D to cut the semicircle. This is very convenient on the drawing board for fairly large ranges of R_2.

The crank OV represents the equivalent impedance of the transformer since we have taken I_1 as unity. To find the current that will run if $E = 100$ V it is only necessary to scale off OV and to divide it into 100.

Reactive Load Variation.

21. The most obvious way of altering the reactive part of the load is to vary L_b, but another way is to introduce capacitance in series with L_b and R_2. We shall first consider the effect of varying L_b while keeping R_2, L and I_1 constant. It is evident that the angle ϕ_2 will increase with increase of L_b since $\phi_2 = \angle L_2\phi/R_2$, and that F will describe part of a semicircle on BD as diameter. Now DV always bears to BF the constant proportion $L\omega : R_2$ and it makes with DH, which is drawn perpendicular to DB in Fig. 154, the angle ϕ_2. Therefore as F describes a semicircle on BD, V describes a semicircle on DH as diameter. If this reasoning is correct the length of DH ought to be independent of L_b and I_2. It is given by similar triangles in Fig. 154 as follows:

$$\mathrm{DH} : \mathrm{DV} = \mathrm{BD} : \mathrm{BF}$$

or
$$\mathrm{DH} = \mathrm{BD} \,.\, \mathrm{DV}/\mathrm{BF} = L\omega I_1 \,.\, L\omega I_2/R_2 I_2$$
$$= L^2\omega^2 I_1/R_2.$$

It is to be noticed that for any given value of L_b the point K has a definite position, but K moves as L_b varies. The geometry of the Figure shows, however, that the three points K, V and H are always on a straight line.

The extreme values of L_b are zero and infinity. When $L_b = 0$ we have $L_2 = L$ and then K is at B. On joining B to H we get the extreme point V_0. Again, when $L_b = \infty$, ϕ_2 is a right angle and V is at D. The necessary primary voltage to drive I_1 through the transformer is always OV, and it is therefore easy to watch the changes in E as V passes from V_0 to D on account of an increase in L_b. As before, the secondary current is proportional to FB and its phase is given by that vector; the primary current is proportional to OA.

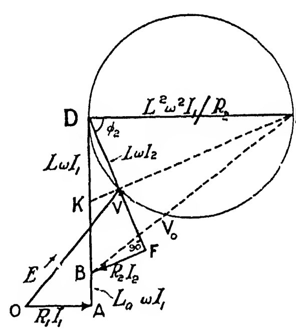


FIG. 154.

Case of Negative Reactive Load.

22. In Fig. 154 the variation of the secondary inductance exterior to the common inductance L has been limited to varia-

tion between zero and infinity, and this range corresponds to motion of V from V_0 to D along the shorter arc. If negative values of L_2 be imagined possible the point V would travel outside the marked range, though still keeping to the circle.

It has already been pointed out that series capacitance is equivalent to negative inductance. A conductor of capacitance C_2 possesses reactance $-1/C_2\omega$ and when in series with a coil of inductance L_2 reduces the reactance to

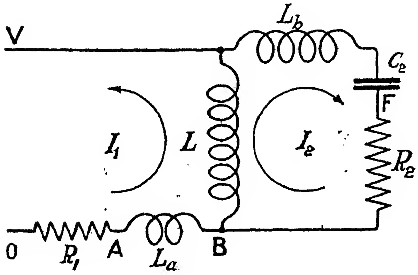


Fig. 155.

$$X_2 = L_2\omega - 1/C_2\omega.$$

For example, it will be possible to attain the point V_0 in Fig. 154 by making $C_2 = 1/L_b\omega^2$ so that all the secondary inductance outside the part common to primary and secondary is annulled. And if C_2 is taken less than this some or all of the common part may be annulled. To annul all the inductance we need only make $C_2 = 1/L_2\omega^2$. If C_2 is made still smaller the total reactance becomes negative. All these possibilities are included in the statement that X_2 may be positive, zero or negative. Throughout all these variations the point V will remain on the circle of diameter DH, as may be seen by substituting X_2 for $L_2\omega$ in the reasoning of § 21. The circuit is given in Fig. 155.

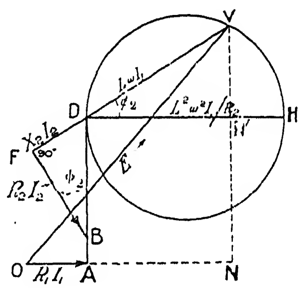


Fig. 156.

23. At every adjustment we shall have the secondary phase angle given by $\tan\phi_2 = X_2/R_2$, and this angle, when calculated from the given data, determines the position of V. When the condenser is so small that X_2 is negative we get a crank diagram like Fig. 156, where ϕ_2 is utilised to determine the triangle of secondary voltages BFD and also the position of V, which gives OV the primary voltage necessary to pass the current I_1 at the adjustment indicated. In this Figure a projector VN′N is drawn to show AN the image resistance and VN′ the image reactance.

The equivalent reactance of the assemblage is VN, which is greater than the primary reactance alone. It may be mentioned in passing that the reactance X' is greatest and least when V is farthest from the diameter DH, and R' is greatest when V is at H, while E is greatest and least when OV passes through the centre of the circle. These particular adjustments lead to many interesting relations between the various quantities, but the writing out of these theorems would take up more space than can be allowed here.

24. In Fig. 156 the secondary inductance and capacitance are lumped together in the term X_2. This has been supposed done by arithmetic, but it can be done graphically. In Fig. 157 DP represents $L_2\omega I_2$ and QB is $I_2/C_2\omega$. The total of these reactive voltages is X_2I_2. The opportunity is taken in this diagram to indicate dotted the circle of variable resistance, X_2 being kept constant. Its diameter DK′ is determined by prolonging HV as in Fig. 154. It will be found that $DK'/DA = L^2\omega^2/L_1\omega X_2$. This may be negative, zero, positive or infinite according to the value of X_2. In Fig. 156 it is negative, in Fig. 154 positive. A wider generalisation still is given in § 27.

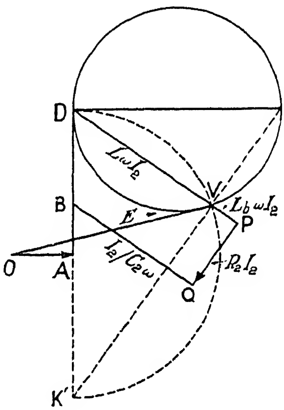


FIG. 157.

25. One important particular case may be given prominence. When V is at H we have $\phi_2 = 0$ which implies that $X_2 = 0$, and therefore $L_2\omega = 1/C_2\omega$. The secondary is here adjusted to resonance, the image reactance is zero and the image resistance is a maximum. Its value is

$$R' = L^2\omega^2/R_2.$$

The power delivery to the secondary is a maximum, and the efficiency $\eta = L^2\omega^2/R_1R_2$ is likewise a maximum. The secondary current is in quadrature with the primary current and the ratio of currents is, numerically,

$$I_2/I_1 = L\omega/R_2,$$

as is seen from the fact that BF must at this adjustment lie along BD.

Capacitance in the Primary.

26. When a condenser of capacitance C_1 is introduced into the primary circuit as indicated in Fig. 158 a vector of magnitude $I_1/C_1\omega$ appears in the crank diagram of sense opposite to the vector of the inductance L_a and may be drawn at any place

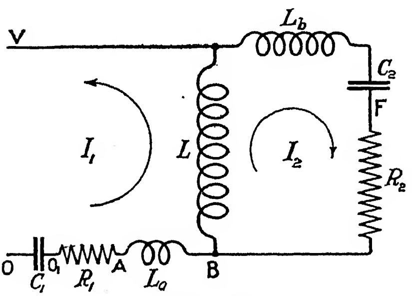


FIG. 158.

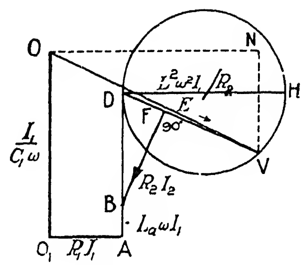


FIG. 159.

convenient for geometrical addition. In Fig. 159 this condenser vector is shown at OO_1. The crank OV represents the necessary primary voltage to give primary current I_1, ON represents the equivalent resistance and NV the equivalent reactance of the assemblage.

As the primary condenser increases in capacitance the length OO_1 decreases. When it is equal to DA and, besides, V at H, we obtain the resonance adjustment of the whole transformer. The primary voltage is then in phase with the primary current. This is seen in Fig. 160. The equivalent resistance is, from the diagram,

$$R' = R_1 + L^2\omega^2/R_2.$$

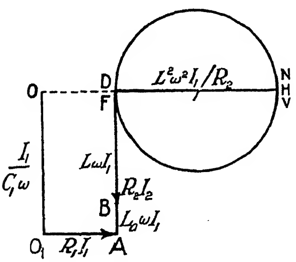


FIG. 160.

The primary and secondary are both behaving as if they possessed no reactance because the condenser in each annuls the inductance with which it is in series.

27. In § 24 is suggested a wider generalisation of the coefficient of coupling. This arises when a condenser is present in the primary circuit. It is made especially clear by supposing for a

time that R_1 is zero. For then the crank diagram becomes that of Fig. 161, in which

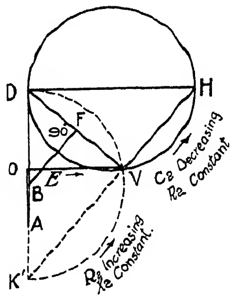


FIG. 161.

$$DK'/DV = DB/DF$$

or $$DK' = L^2\omega^2 I_1/X_2,$$

and $$DO = X_1 I_1,$$

and therefore $DK'/DO = L^2\omega^2/X_1X_2$.

This may be styled the coefficient of coupling between the *circuit reactances* as distinct from that between the *circuit inductances*. Since both X_1 and X_2 may have any value positive or negative, the position of K′ has no limits.

28. A curious particular instance arises when the constants of the circuits are such that the point O bisects the diameter DK′ of the resistance semicircle. For in that case we have Fig. 162 wherein the applied voltage E needed to give the constant primary current I_1 is itself constant for all values of the load R_2, other things being unchanged. For this to happen we must have the generalised coefficient of coupling equal to 2, whence

$$L^2\omega^2 = 2X_1X_2.$$

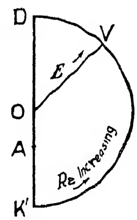


FIG. 162.

Non-reactive Adjustments.

29. The resonance condition just discussed is not the only adjustment that will make the assemblage behave as a mere resistance, that is, make the power factor unity. For in Fig. 159 we need only draw a line through O parallel to O_1A and, in general, this will cut the circle and determine two possible positions for V. Provided that the circle is intersected by the parallel to O_1A two values of C_2 can be found in general to give unity power factor. But the two values coincide if it happens that the line through O is tangent to the circle. It is plain that if we are allowed to choose any value for the primary condenser we can always obtain a non-reactive adjustment. But if there is no primary condenser at all the only chance of obtaining the non-reactive adjustment is by making R_2 small enough; for by

reducing R_2 we increase the diameter of the circle of Fig. 154 and so make it touch or cut twice the line OA prolonged.

30. The case of tangency is shown in Fig. 163. Clearly AD must equal the radius of the circle; whence we get

$$L_1\omega I_1 = \tfrac{1}{2}L^2\omega^2 I_1/R_2$$

or
$$R_2 = L^2\omega/2L_1.$$

Besides having the circle of the correct magnitude we must have V at the proper place and to ensure this we must have $\phi_2 = 45$ degrees. This means that $X_2 = R_2$. Whence

$$L_2\omega - 1/C\omega_2 = R_2$$

or
$$C_2\omega = 1/(L_2\omega - R_2).$$

For example the transformer of § 20, in order to fill these conditions, must have a secondary resistance

$$R_2 = 16 \times 314/2 \times 5$$
$$= 503\ \Omega$$

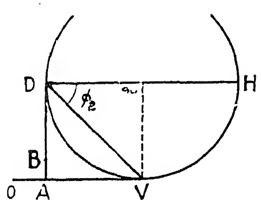


FIG. 163.

and also a secondary condenser of capacitance

$$C_2 = 1 \div (4{\cdot}02 \times 314 - 503)\,314$$
$$= 1 \div (760 \times 314)$$
$$= 4{\cdot}2\ \mu\text{F}.$$

It will be found that if R_2 be eliminated between the two equations above, the condition

$$L_2C_2\omega^2(1 - \tfrac{1}{2}k^2) = 1$$

is obtained, when k is the coupling coefficient of the inductances. This equation may be regarded as one to give ω in terms of the constants of the circuit.

31. As shown by Fig. 164, if R_2 is less than that critical value which makes the circle tangent to OA prolonged we obtain for V two positions in which the assemblage is non-reactive. The values of C_2 necessary to make the equivalent reactance zero are obtained by solving an equation expressing that the projection of DV on AD is equal to AD. This equation is DV $\sin\phi_2 =$ AD. Now DV $= L\omega I_2$, $\sin\phi_2 = X_2/Z_2$ and AD $= L_1\omega I_1$, therefore

$$L\omega I_2 X_2/Z_2 = L_1\omega I_1$$

or, using the equation
$$I_2/I_1 = L\omega/Z_2,$$

$$L^2\omega X_2 = L_1(X_2{}^2 + R_2{}^2).$$

This is a quadratic equation for X_2. When X_2 is determined the

value of C_2 is immediately deducible since L_2 is supposed given. The solving of the quadratic will be left to the reader. For most practical purposes it is best to solve for C_2 graphically. All that has to be done is to set off DP to represent $L_2\omega I_2$ to the same scale as DV represents $L\omega I_2$, draw PQ and QB perpendicular and parallel, and scale off QB, which represents $I_2/C_2\omega$. The construction must be repeated at V′. (Fig. 164.)

It must be remarked that the quadratic equation for X_2 leads on writing out in full to a quadratic in ω^2; proving that for a given circuit there are two frequencies giving the non-reactive phenomenon, provided that R_2 is less than the critical value.

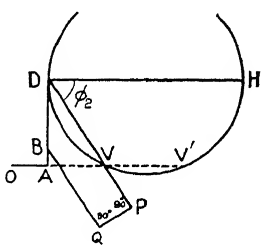


FIG. 164.

32. An important application of the above arises when $L_1 = L_2 = L$, that is to say when the only inductance in primary and secondary is the coupling coil. Then the quadratic for X_2 becomes

$$L\omega X_2 = X_2^2 + R_2^2.$$

But

$$L\omega = X_2 + 1/C_2\omega$$

therefore

$$X_2/C_2\omega = R_2^2.$$

This gives

$$\frac{L}{C_2} - \frac{1}{C_2^2\omega^2} = R_2^2$$

or

$$1/\omega^2 = LC_2 - C_2^2R_2^2$$

which is a quadratic equation for C_2; or which may be regarded as an equation for the frequency at which the assemblage is non-reactive, all the constants being given. The graphical solution of this particular case is obtained very simply, for A becomes coincident with B and P with V in Fig. 164.

It should be noticed that if, in any of these non-reactive assemblages, the primary resistance R_1 is negligible, then the assemblage offers no resistance to direct current though it offers the resistance R' to alternating current.

Direct Electric Coupling

33. A typical circuit is described by Fig. 165 and is the analogue of Fig. 150 which exhibited typical direct magnetic

coupling. We shall indicate the total primary capacitance by C_1 where

$$1/C_1 = 1/C_a + 1/C$$

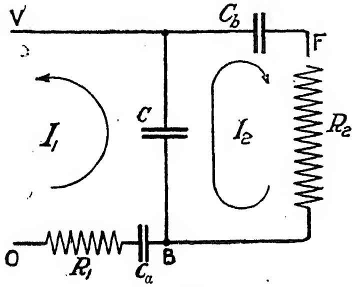


FIG. 165.

and the total secondary capacitance by C_2 where

$$1/C_2 = 1/C_b + 1/C.$$

The crank diagram appears in Fig. 166 and is constructed as follows: — Draw OA to represent the primary ohmic potential drop, set off AB at right angles to represent the potential drop in C_a to the same scale, and then BD in the same line to represent the drop in C. This last potential difference, being introduced by the primary current into the secondary circuit, may be said to produce the secondary current and so give rise to the vector triangle DFB, which is constructed by making at B the angle $\phi_2 = \angle\, 1/C_2\omega R_2$, and dropping from D a perpendicular to define the point F. Then DF is $I_2/C_2\omega$ and FB is R_2I_2. From D mark off DV the same proportion of DF as $1/C$ is of $1/C_2$, and join OV. Then DV represents $I_2/C\omega$ and VF equals $I_2/C_b\omega$. Thus DV is the reaction of the secondary circuit upon the primary, and this adds geometrically to the cranks OA and AD to give the applied emf in the primary. That is OV $= E$, the emf needed to force the constant current I_1 through the assemblage. In fact, if I_1 is made unit current, OV is the impedance of the assemblage and AOV its phase angle. By projecting V upon OA and AD, that is to say, by drawing the perpendiculars VN and VN′, we obtain the voltage corresponding to the image resistance R' at AN and that corresponding to the image reactance at N′D. The equivalent ohmic drop and reactive drop are seen at ON and NV.

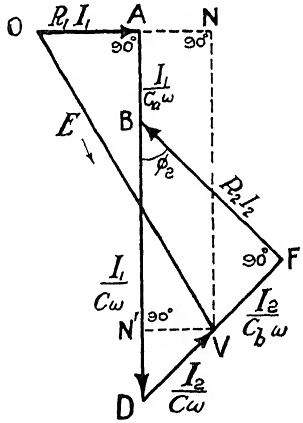


FIG. 166.

34. It will be seen on comparison with §17 that there is a very close analogy between the electric and magnetic problems. The crank diagram of Fig. 166 need only be turned over about OA as axis to become Fig. 151, provided that $L_a\omega$ be made equal to $1/C_a\omega$, $L\omega$ equal to $1/C\omega$ and $L_b\omega = 1/C_b\omega$. All this is included in the statement that, mathematically speaking, we may substitute $-1/C\omega^2$ for L, and so forth.

35. The formulæ and constructions for the ratio of currents, the coupling coefficient, and the variation of load, follow immediately by the application of this substitution to preceding paragraphs.

The ratio of currents is

$$I_2/I_1 = 1/C\omega Z_2,$$

neglecting phasal relations, where

$$Z_2 = \{R_2^2 + (1/C_2\omega)^2\}^{\frac{1}{2}}.$$

The coupling coefficient in Fig. 165 is

$$k^2 = C_1C_2/C^2.$$

Example of Varied Frequency. Electric Absorption.

36. A case of great interest is that of Fig. 167, where two condensers, C, K, connected in series have a resistance R in series with them and a resistance S in shunt with K. The vector diagram is given in Fig. 168 for two frequencies, one double the other. In the lower frequency diagram the condenser P.D's are doubled, the drop in R is unaltered (because I is taken of unaltered value), and the shunt triangle has its phase angle practically doubled. Evidently there is but small change in the angle between the emf vector OV and the current vector OI when the electrical constants are taken of about the magnitude chosen here; for the diminution of frequency in passing from one diagram to the other moves V to the right while it also moves it down the page relative to O. Thus a combination of condensers and resistances is easily hit upon to have a constant power factor over a large range and frequency. It has been found by experiment that a single condenser with an ordinary solid dielectric

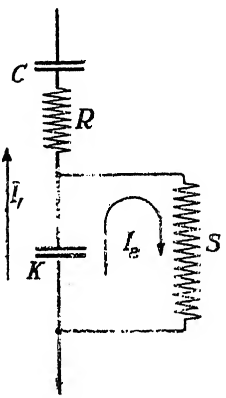


FIG. 167.

between its plates possesses this property of unvarying power factor over a wide range of frequencies. This leads to the suggestion that such a condenser can be represented by an arrangement like that of Fig. 167; or, doubtless more closely still, by an assemblage of such arrangements.

An assemblage such as this would display some of those properties included under the terms "electric absorption" and "soaking in." For on applying a voltage to this circuit for a long time, always under the supposition that S is a very high resistance, current would flow in slowly to a greater amount than would be accounted for by the nominal value of the capacity for relatively quick chargings and dischargings; while on discharging and insulating the terminals after slowly charging, some of the trapped electricity on the plates connected to R would leak across S and establish a slight difference of potential, leading to a further slight discharge on joining the terminals.

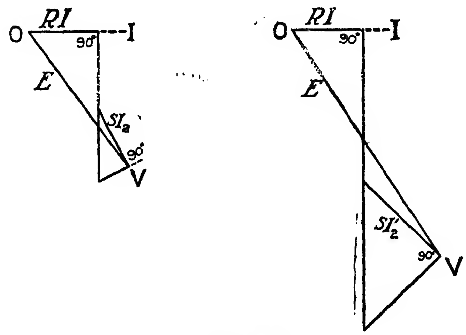


FIG. 168.

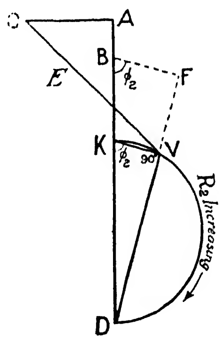


FIG. 169.

Variation of Ohmic Load.

37. The construction for a varying resistance load is given in Fig. 169. Here VK is perpendicular to DF and DK/DA $= k^2$. As R_2 increases the angle ϕ_2 decreases and the point V moves towards D.

The numerical example of § 20 may be reworked by the reader with L_a replaced by a condenser of 10·13 μF, L_b by one of 506·5 μF, and L by one of 2·53 μF.

Variation of Reactive Load.

38. Reactive load variations require the crank diagram of 170. The portion of the circle between D and V_0 is described

clockwise by V when the load capacitance C_b of Fig. 165 is increased from zero to infinity. To go past V_0 inductive load must be introduced. Let this be of amount L_2 in series with R_2. As L_2 increases V moves from V_0 to H, where the resonance

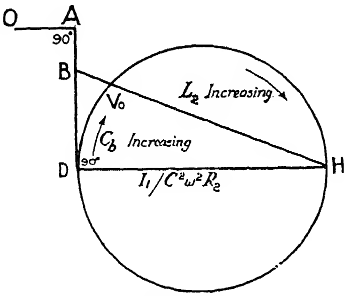


FIG. 170.

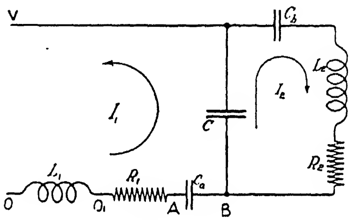


FIG. 171.

value is attained, namely, $L_2 = 1/C_2\omega_2$. If L_2 is increased beyond this value V describes the semicircle clockwise from H to D, at which latter point L_2 becomes infinite.

When inductance is introduced into the primary we obtain the circuit of Fig. 171 and the crank diagram of Fig. 172. Here the resistance semicircle is drawn in as well as the variable reactance circle. It is to be observed that for every different value of resistance load R_2 there is a different circle DH, and for every value of secondary reactance there is a different semicircle DK′. The adjustment V′ in Fig. 172 indicates an over-inductive secondary with semicircle on diameter DK″ laid in opposite direction from the semicircle on DK′.

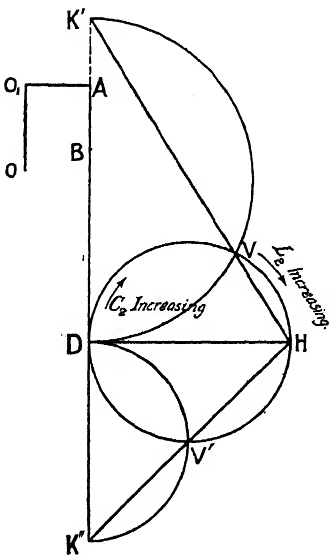


FIG. 172.

The curious particular case discussed in § 28 can be duplicated by a condenser coupling. The reader will be able to examine the properties of the condenser assemblage by aid of that paragraph.

Non-Reactive Adjustments.

39. Non-reactive adjustments for the condenser transformer follow immediately by translating §§ 29 to 32. The only circuit that needs special mention here is one much used in connection with ionic tubes which is sketched in Fig. 173. The crank diagram of Fig. 174 shows the functions of each part. The applied voltage E is shown at OV in phase with the current I_1. The algebraic condition for this is obtained from § 31 by sub-

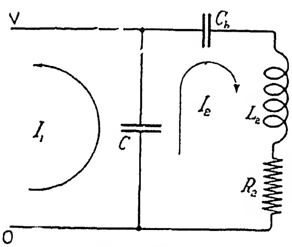


FIG. 173.

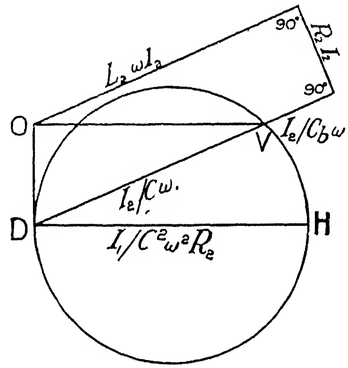


FIG. 174.

stituting $-1/C\omega^2$ for L and $-1/L_2\omega^2$ for C_2, or it may, of course, be obtained direct from Fig. 174. In either way we have

$$\omega^2 = \frac{1}{L_2C} - \frac{R_2^{\,2}}{L_2^{\,2}}$$

as the relation between the frequency and the constants in the non-reactive condition. Obviously R_2^2 must be less than L_2/C for the adjustment to be possible. With this adjustment we have OV $= R'I_1$

and therefore $\quad (R'I_1)^2 = (I_2^{\,2} - I_1^{\,2})/C'^2\omega^2$

or $\quad R'^2 = \{(I_2/I_1)^2 - 1\}/C'^2\omega^2.$

Since, always, $\quad (I_2/I_1)^2R_2 = R'$

we have $\quad R'^2C'^2\omega^2 = R'/R_2 - 1.$

The solution of this quadratic for R', the image resistance, is given by

$$2R'R_2 = 1 \pm (1 - 4R_2^{\,2}C'^2\omega^2)^{\frac{1}{2}}.$$

Mixed Reactive Coupling.

40. An autotransformer may be made up with both inductance and capacitance in the common part, as indicated in Fig. 175. Since the P.D. across L and C in series is less than that across one of them we should expect the secondary currents to be in general smaller than if either part were acting alone as the coupling agency. Indeed it is to be expected that when $L\omega$ is made equal to $1/C\omega$ the coupling action will be zero.

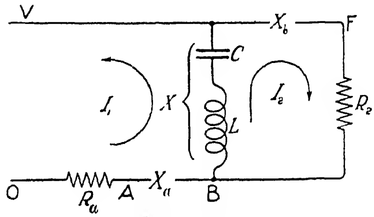


FIG. 175.

The crank diagram is readily constructed for any given data. The typical diagram given in Fig. 176 is constructed on the supposition that all the reactances marked in Fig. 175 are positive. The reasoning concerning the various cranks and the circular loci is precisely as in §§ 10, 19, 37, etc. The ratio DK′/DB, the coefficient of reactance coupling, is equal to $X/(X_1X_2)^{\frac{1}{2}}$. Since there are here three quantities that can range through positive and negative values there are a great many particular cases which yield diagrams of widely different appearance.

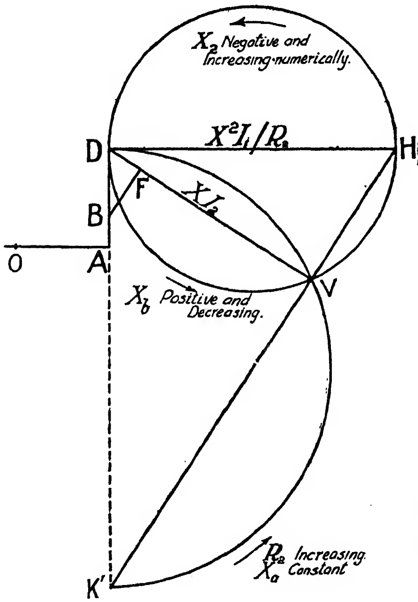


FIG. 176.

Mixed Coupling.—General Case.

41. The most general form of direct coupling arises when the common branch of an autotransformer possesses both resistance and reactance. Fig. 177 gives the circuit diagram

and Fig. 178 the crank diagram. In the latter diagram the reactances X_a and X are supposed to be positive and X_b is taken negative. In working out the crank diagram we start with OA to represent $R_a I_1$, and draw AB at right angles to represent $X_a I_1$. Then the vectors for RI_1 and XI_1 follow naturally; their resultant is marked in the Figure as BD $= ZI_1$.

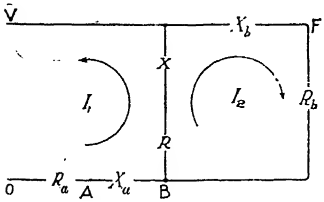


FIG. 177.

This vector represents the magnitude and phase of the P.D. introduced into the secondary circuit by the flow of primary current through the common impedance. On BD therefore we construct the secondary crank diagram DPQB as follows: Make the angle DBF equal to $\phi_2 = \angle X_2/R_2$, and draw DF perpendicular to BF. Make

$$\text{DP} = (X/X_2)\ \text{DF},$$

draw PQ parallel to FB and BQ parallel to DP. Then PQ is $R_2 I_2$, DP is XI_2, QB is $X_b I_2$. (In the Figure X_b is taken negative.) In order to determine the reaction of the secondary current upon the primary circuit make the angle PDV = B′DB and join DV. Then DV represents ZI_2 the secondary reaction, therefore OV (not drawn in the Figure) is E the primary applied emf.

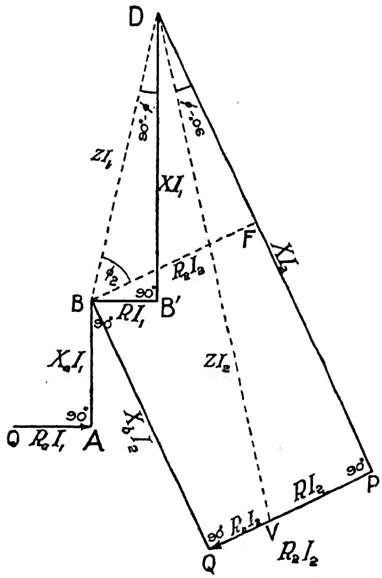


FIG. 178.

42. The reasoning of § 19 may be easily adapted to prove that the locus of V is a part of a circle when R_2 is varied with X_2 constant, and that the locus is another circle when X_2 is varied and R_2 kept constant. In order to construct these loci draw DH so as to make the angle BDH

a right angle, and draw DK′ so that K′DB = $\angle X_2/R$. Through V draw a perpendicular to DV to determine K' and H. On DK′ as diameter describe a semicircle—this is the locus for variation of ohmic load. On DH as diameter describe a circle—this is the locus for variation of reactive load.

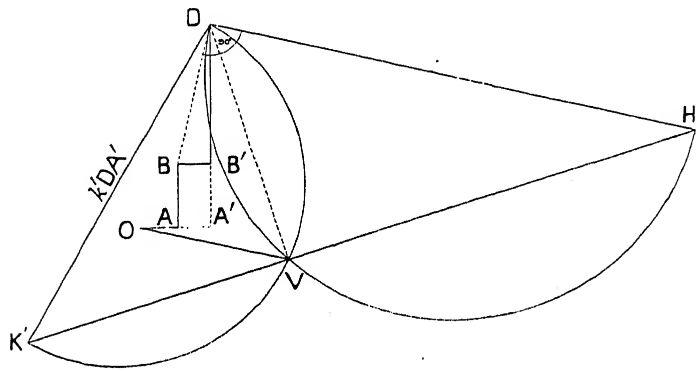


FIG. 179.

By adaptation of the methods already used in the simpler cases we find, on prolonging OA and DB′ to meet at A′, that

$$DK'/DA' = Z^2/X_1X_2 = k^2 \text{ and } DH = (Z^2/R_2)I_1.$$

The circles are not drawn in Fig. 178, but appear in Fig. 179, which is constructed with the fewest possible lines, as follows :—

$$OA = R_aI_1,\ AB = X_aI_1,$$
$$BB' = RI_1,\ B'D = XI_1,$$

join DB, make the angle BDK′= $\angle X_2/R$, draw DH perpendicular to DB, make DH = Z^2I_1/R_2 and DK′ = k^2DA'. Draw circles on DH and DK′ as diameters, intersecting at V. Join DV, HK′ and OV.

INDIRECT MAGNETIC COUPLING

43. If the coil common to both circuits in Fig. 150 and marked L be imagined to have its wire split and separated infinitesimally as suggested by Fig. 180 the secondary circuit will become wholly detached from the primary. The magnetic flux due to unit current flowing in the primary and traversing one half of L will send through the other half of L the same magnetic flux as formerly, and consequently an alternating current in the

primary will produce in the secondary an emf equal to and opposite from the potential drop produced at the primary terminals of L. As already explained the total of the flux linkages with either circuit due to unit current in the other is called the mutual inductance between the coils and that total is indicated by M. It is obvious that in the case of the split coil $M = L$. Since the primary and secondary conditions are quite unchanged the crank diagram will be precisely the same for this indirectly coupled pair of circuits as for the original directly coupled pair.

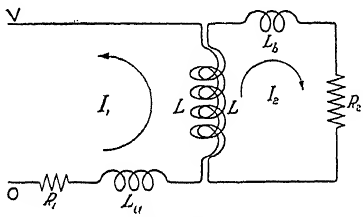


FIG. 180.

Imagine now the primary and secondary moved apart. The inductance of the primary remains of the value $L_1 = L_a + L$, that of the secondary remains $L_2 = L_b + L$, but the flux linkages produced in either circuit by unit current in the other are less numerous, and therefore the crank diagram will be changed notably in respect of the vectors concerned with the action and reaction between the circuits. The action of the primary on the secondary is reduced to the fraction M/L_1 of the primary inductive drop. In Fig. 181 BD represents $M\omega I_1$, the action of the primary on the secondary under these changed conditions, that is M/L_1 of AD. The triangle of secondary voltages BFD is constructed by making the angle ϕ_2 at B equal to the secondary phase angle and dropping a perpendicular DF on its bounding line. This determines DF as representing $L_2\omega I_2$ and FB as $R_2 I_2$. The reaction from the secondary upon the primary is still strictly in phase with the self-inductive

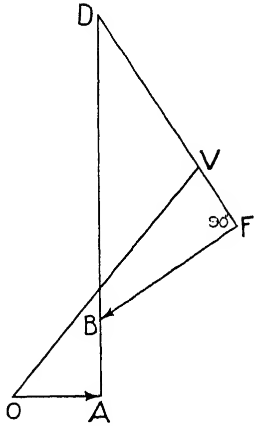


FIG. 181.

drop $L_2\omega I_2$ because it is due to part of the same body of flux. Therefore mark off $DV = M/L_2$ of DF so that $DV = M\omega I_2$. Then OV is the primary emf. The diagram may be read as follows: The ohmic drop OA with the primary inductive drop

AD and the inductive drop due to linkage with the secondary flux DV make up the applied emf OV. Or alternatively it may be read: The ohmic drop OA plus the inductive drop AD is made up by the primary emf OV assisted by the reaction VD from the secondary.

Algebraic Solution.

44. The symbolic statement of these facts is obtained by writing down that the primary and secondary currents are each equal to the total emf in the respective circuit divided by the resistance. By total emf is meant that applied by a source such as an alternator added geometrically to the emf induced according to Faraday's law by the decrease of flux linkages. Thus since $L_1 i_1 + M i_2$ is the primary flux linkage and $L_2 i_2 + M i_1$ is the secondary flux linkage we have

$$i_1 = \{e - D(L_1 i_1 + M i_2)\} \div R_1$$

and

$$i_2 = -D(L_2 i_2 + M i_1) \div R_2$$

as the two equations for the two currents to be determined. Adopting the potential drop point of view—which merely means transferring all terms except those due to external sources to the lefthand sides of the equations—we obtain

$$(R_1 + L_1 D) i_1 + MD i_2 = e$$
$$(R_2 + L_2 D) i_2 + MD i_1 = 0.$$

In order to solve for i_1 operate on the first equation by $R_2 + L_2 D$ and on the second by MD and subtract. We obtain

$$i_1 = \frac{(R_2 + L_2 D)e}{(R_1 + L_1 D)(R_2 + L_2 D) - M^2 D^2}$$
$$i_2 = -\frac{MD i_1}{R_2 + L_2 D}.$$

The latter equation is easily reduced and yields

$$i_2 = \frac{(M\omega, -\tfrac{1}{2}\pi)}{(Z_2, \phi_2)} i_1$$
$$= \left(\frac{M\omega}{Z_2}, -\tfrac{1}{2}\pi - \phi_2\right) i_1$$

which gives the current ratio as

$$I_2 / I_1 = M\omega / Z_2$$

and shows that i_2 lags behind i_1 by the angle $\frac{1}{2}\pi + \phi_2$. Here

$$Z_2^2 = R_2^2 + L_2^2 \omega^2$$
$$\tan \phi_2 = L_2 \omega / R_2.$$

The former of the two equations may be reduced in the manner of III, § 42 and the equivalent resistance and reactance fully expressed, but it is briefer, and better for numerical work, to write

$$i_1 = \frac{(Z_2, \phi_2)e}{(Z_0, \phi_0)} = \frac{e}{(Z_0/Z_2, \phi_0 - \phi_2)}$$

where Z_0 and ϕ_0 are obtained by multiplying out the denominator and substituting $-\omega^2$ for D^2. We thus obtain

$$Z_0^2 = \{R_1R_2 - (L_1L_2 - M^2)\omega^2\}^2 + \{R_1L_2 + R_2L_1\}^2$$

$$\tan\phi_0 = \{R_1L_2 + R_2L_1\} \div \{R_1R_2 - (L_1L_2 - M^2)\omega^2\}.$$

It is plain that the equivalent impedance of the whole transformer is Z_0/Z_2, and its phase angle $\phi_0 - \phi_2$; these are easily calculated from numerical data by aid of the above results. It will, however, be found that for many practical purposes the drawing board method is better than this arithmetical one.

Analysis of the General Case.

45. The algebraic solution of the general case of two inductively coupled circuits will be added very briefly. Let R_1, X_1 be the resistance and reactance of the primary, and R_2, X_2 of the secondary, of the coupled circuits of Fig. 182. Then the potential drop equations for primary and secondary are, by § 44.

$$(R_1 + X_1D/\omega)i_1 + MDi_2 = e$$

$$MDi_1 + (R_2 + X_2D/\omega)i_2 = 0.$$

On eliminating i_2 we obtain

$$\frac{e}{i_1} = \frac{R_1R_2 - (X_1X_2 - M^2\omega^2) + (R_1X_2 + R_2X_1)D/\omega}{R_2 + X_2D/\omega}.$$

The secondary current is expressed in terms of the primary current by the second of the original equations without further labour beyond "rationalising" to remove D.

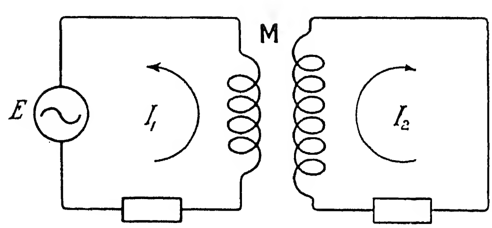


FIG. 182.

Non-Reactive Adjustments.

46. The result of § 45 shows that the assemblage is non-reactive if

$$\frac{R_1R_2 - X_1X_2 + M^2\omega^2}{R_2} = \frac{R_1X_2 + R_2X_1}{X_2} = Z, \text{ say,}$$

or $\quad (M^2\omega^2 - X_1X_2)/R_2 = R_2(X_1/X_2) = Z - R_1.$

Here Z is the non-reactive impedance, that is, the equivalent resistance of the assemblage.

These results show that in order to make a pair of coupled circuits non-reactive it is necessary, first, to have X_1 and X_2 of the same sign, and, second, to have $M^2\omega^2 > X_1X_2$. This latter condition necessitates having a condenser in either the primary or the secondary. The former condition implies that both primary and secondary must have simultaneously too little or simultaneously too much capacitance for resonance.

47. The equation above, which may be written

$$X_1X_2^2 = M^2\omega^2X_2 - R_2^2X_1,$$

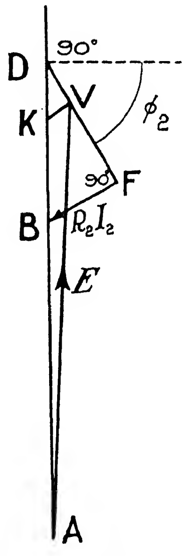


FIG. 183.

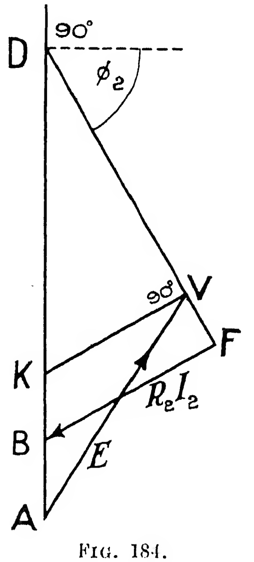


FIG. 184.

is the equation for the non-reactive frequencies. It is a cubic in ω^2, indicating that there might be three possible non-reactive frequencies. The equation reduces to a quadratic in a number of cases, among which may be mentioned: either circuit being without a condenser, or the circuits being in resonance with each other, or the coupling coefficient being unity.

Crank Diagrams with Mutual Inductance Varied.

48. Returning to the crank diagram let us start in Fig. 180 with the circuits well separated, that is with M small, and let us imagine that the windings can be gradually brought together till

the turns severally touch and finally coalesce to form the auto-transformer. Let us suppose that $L_a = 0$, $L_b = 0$ and $R_1 = 0$. Suppose, too, that the primary current, and therefore $L\omega I_1$, is kept constant as the windings are moved together and that E is suitably varied for the purpose. At first, with the windings well separated, we have Fig. 183. Then as the mutual linkage of flux increases, the E vector swings clockwise and shortens as in Fig. 184 until at the moment of coalescence of the windings it becomes perpendicular to the secondary inductive vector. At the same time the secondary ohmic drop vector draws into coincidence with E, or in other words, I_2 increases until the ohmic drop is numerically equal to the applied emf. This is indicated in Fig. 185. In this limit the angle between E and I_1 is the least of all the values assumed, that is to say, the power factor is the greatest. It will be noted that the sense of the ohmic drop vector is opposite that of the primary voltage vector.

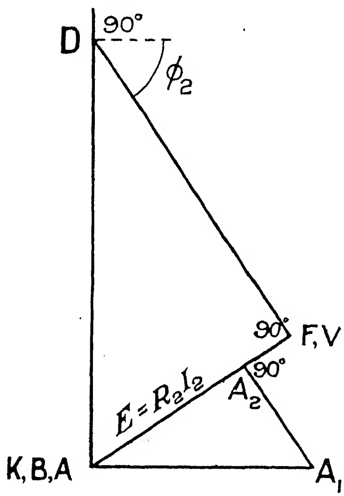


Fig. 185.

This case is approximately that of the "commercial" transformer on lighting load; for in the design of this transformer the magnetic leakage is made very small. Taking due account of the relative directions of the currents as defined by the cyclic notation of Fig. 180, we see that the current produced in the secondary is given by

$$I_2 = E/R_2,$$

just as if the secondary resistance were in the primary circuit and the transformer non-existent. But the current actually running in the primary of the transformer in this limiting case of perfect coupling is not in phase with the applied primary emf nor equal to the secondary current. In fact the currents are in the ratio of BD to DV. Therefore in Fig. 185 if AA_2 represents the secondary current to scale, AA_1 represents the primary current to the same scale; and AA_1 may be regarded as made up of a power component equal to the secondary current and a wattless component

A_2A_1. This mode of resolving is explained in III, § 31. The wattless component is commonly called the magnetising current of the transformer.

Remarks on Mutual Inductance.

49. So far the reasoning in the most general case has supposed that the coils mutually affecting each other were equal and that the remainder of the inductance in either circuit was distinct. There is no necessity for this limitation. The coils $L_{\prime\prime}$ and L may be concentrated into one coil of inductance $L_1 = L_{\prime\prime} + L$, and the secondary may be treated similarly, without invalidating the diagrams already constructed. Of course if the original value of M is to be maintained, the larger coils must be readjusted in position to bring this result about.

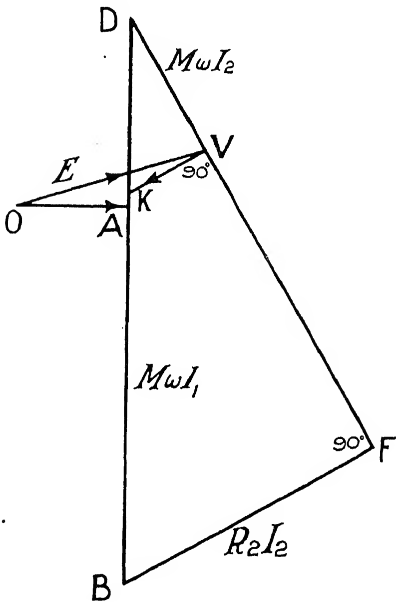


FIG. 186.

When very unequal coils are put into mutual relation the value of the mutual inductance may be larger than the self-inductance of the smaller coil. It is, of course, necessary, as provided in II, § 64, that M^2 shall be less than L_1L_2. As an example let $L_1 = 10^6$, $L_2 = 10^4$ and $M = 5 \times 10^4$. Then, clearly, the above statements are justified. The consequence of this is that in the general case the diagrams may take either of the shapes shown in Fig. 186 and Fig. 187.

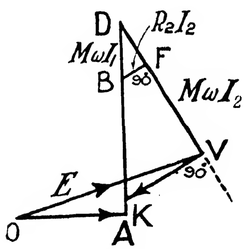


FIG. 187.

In these Figures the ratio of the secondary to the primary current is DV/DB, and the ratio of the secondary resistance terminal voltage to the primary is given by FB/OV. Therefore in the two transformers represented by these Figures the one with $M > L_1$ steps current down and voltage up, while

that with $M < L_1$ does the reverse. The former has more inductance in the secondary than in the primary, the latter more in the primary than in the secondary. The coefficient of coupling is the same in both transformers, as is shown by the position of K.

Secondary Resistance Variable.

50. The variation of secondary resistance directly affects the value of ϕ_2, the secondary phase angle, which decreases as R_2 increases; but it has no influence on the position of K relative to D. Hence by the reasoning of § 19 we see that V describes a semicircle on KD as diameter. This is exhibited in Fig. 189.

When R_2 is zero, V is at K; when R_2 is infinite, V is at D. In other words, K is the short-circuit point and D is the open-circuit point. At the latter point the secondary might as well not exist, and the voltage E is exactly that required to drive I_1 through the primary winding. Any general position such as V is dictated by the angle DKV, whose tangent is secondary reactance divided by secondary resistance. The lengths of VK and of OA enable the amplitudes of the primary and secondary currents to be estimated and their directions give the phases. The secondary resistance terminal voltage is given by multiplying VK by (L_2/M).

Numerical Example.

51. Consider a numerical example. Let the antenna of Fig. 188 be excited by continuous waves and let a steady current of 1 ampere be running in it. It is shown coupled to an aperiodic circuit comprising inductance L_2 and a thermal instrument of resistance R_2. Let the given data be $L_1 = 4$ mH, $L_2 = 10$ mH, $M = 5$ mH, $R_1 = 20\ \Omega$, $R_2 = 20{,}000\ \Omega$, $\omega = 10^6$ radians per second, and let it be required to find the secondary current and the necessary antenna voltage. First note that $L_2\omega = 10^{-2} \times 10^6 = 10^4$ and that

$$L_2\omega/R_2 = 0{\cdot}5.$$

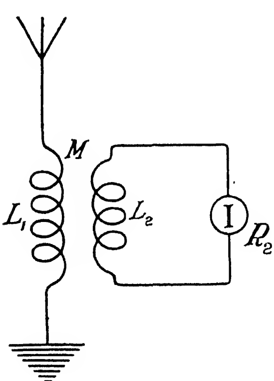


FIG. 188.

This is the tangent of 26°34′. The angle 90° − 26°34′ is set off in Fig. 189 as the angle KDV. Next mark off DA on any convenient scale to represent $L_1\omega I_1 = 4 \times 10^{-3} \times 10^6 = 4{,}000$ volts (the

primary current I_1 being given 1 ampere); and on DA determine K so that DK/DA $= M^2/L_1L_2 = 25/40 = 5/8$. On DK draw a semicircle cutting DV in V. Lastly mark AO $= R_1I_1 = 20$ on the scale adopted for DA, and join OV, VK. Arrows have been added in the Figure to show the directions of the E, I_1, I_2 vectors. The secondary terminal voltage is given by VK multiplied by L_2/M, which is 2. The value of VK is easily calculated; it is given by

$$\begin{aligned} \text{VK} &= \text{DK} \sin 63°26' \\ &= 2{,}500 \times 0{\cdot}8945 \\ &= 2{,}236 \text{ volts.} \end{aligned}$$

Therefore the secondary resistance terminal voltage is 4,472 volts. From this we get the secondary current

$$\begin{aligned} I_2 &= 4{,}472 \div R_2 \\ &= 0{\cdot}2236 \text{ ampere.} \end{aligned}$$

All these numbers are reckoned per ampere of primary current.

52. In order to determine the primary terminal voltage, we must calculate the length of OV. This is tedious, unless a drawing-board measurement will serve; but an approximate value is easily obtained arithmetically by neglecting OA. The problem then reduces to solving the triangle AKV. The most fundamental method is to write down

$$\text{AV}^2 = \text{AK}^2 + \text{KV}^2 + 2\text{AK} \,.\, \text{KV} \cos \angle \text{DKV}.$$

Noting that the cosine of DKV is equal to the sine of KDV we get, to slide rule accuracy, $E = 3{,}640$ volts. The ratio of voltage transformation is thus 1·23 up, while the ratio of current transformation is 4·47 down. The above value of E, it must be emphasised, is reckoned per ampere of primary current.

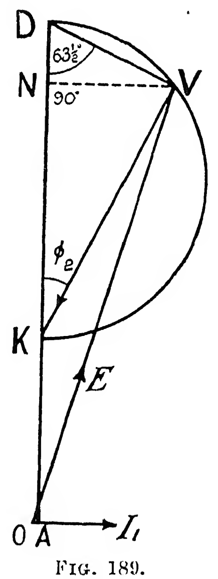


FIG. 189.

In the present example the coupling is close; when it is very small the semicircle becomes minute and useless for forming accurate estimates of the various variables. In such cases formulæ, not diagrams, must be used. We shall use the diagrams later for deducing the necessary formulæ.

Efficiency Considerations.

53. The semicircle indicates how V moves as R_2 is varied, all other things except E and I_2 being invariable. If the resistance is increased, V moves towards D; if decreased, towards K, and, as shown by the change of length of DV, the secondary current varies oppositely in the respective cases. These changes raise the question, very important in practice: What is the value of R_2 giving the maximum rate of absorption of energy (I_1 supposed constant)? This is answered by noticing that the total power absorbed by the antenna and its secondary is proportional to the projection of the vector E on the direction of I_1, which is OA + NV, where NV is perpendicular to AD. Now OA is proportional to the power spent in R_1 and therefore NV must be proportional to the power delivered to R_2. Obviously NV is greatest when N is at the centre of the circle and the angle KDV = 45 degrees. For this to happen in the example we must take $R_2/L_2\omega = 1$; that is, $R_2 = L_2\omega = 10{,}000$ ohms. The various vectors in the diagram are easily determined in this position.

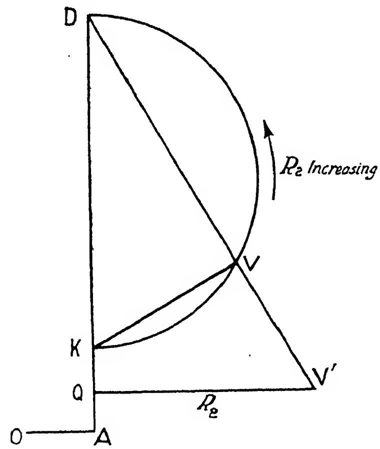


FIG. 190.

Approximately the primary voltage is 2,900, the secondary 3,535, and the secondary current 0·3535. The voltage ratio is therefore 1·22 up, and the current ratio is 2·83 down.

54. In the last paragraph I_1 has been kept constant. Suppose on the other hand that the primary voltage is kept constant, as would more nearly be the case if waves of constant amplitude were sweeping past the antenna. Then we must study the variations as V moves, not of NV itself, but of NV/OV, which is the power absorbed per volt on the antenna. An accurate result can be obtained by plotting from the diagram a curve of this ratio for various values of R_2 on an R_2 base line, and observing the maximum ordinate; but as OA is very small in the present example we may obtain an approximate result by neglecting

OA. Then the power per volt is proportional to NV/AV, or the sine of the angle NAV. This is greatest when the angle is greatest, which is when OV is a tangent to the circle. At this moment the angle KDV is half the complement of the angle DAV, and this latter angle has its sine equal to $2{\cdot}5/(3+2{\cdot}5)$, since K divides DA in the ratio 5 to 3. Hence $R_2 = L_2\omega \times 0{\cdot}6124 = 6{,}124$ ohms for maximum absorption of energy from waves of constant amplitude.

55. In actual work with permanent circuits it is convenient to be able to determine quickly the value of R_2 corresponding to any assigned position of V. A way of doing this on a diagram once drawn has been given in § 20 and is again shown in Fig. 190. Take any position of V for which R_2 is known by calculation from the angle KDV and find on DV a point V′, so that a perpendicular V′Q represents that value on a convenient scale. Now divide QV′ uniformly on that scale. In future R_2 can be read off for any position of V by merely prolonging DV to cut QV′.

Equivalent Resistance and Reactance.

56. These quantities have been defined in §§ 14, 15 for direct coupling in terms easily transferred to the problem of indirect coupling. The formulæ obtained there can be modified to suit the new case, but they are easily deduced anew from Fig. 191. The apparent increase over the primary resistance is AN/I_1, and the apparent decrease from the primary reactance is $N'D/I_1$. But we have seen that $AN = R_2 I_2^2/I_1$ and therefore the apparent increase of resistance is $R_2(I_2/I_1)^2$. This is, in fact, obvious more directly; for if $\frac{1}{2}R_2 I_2^2$, the power spent in the secondary, has to be represented as spent in an image resistance r by a current I_1, we must have $rI_1^2 = R_2 I_2^2$, whence $r = R_2(I_2/I_1)^2$. The total effective resistance between the primary terminals is therefore

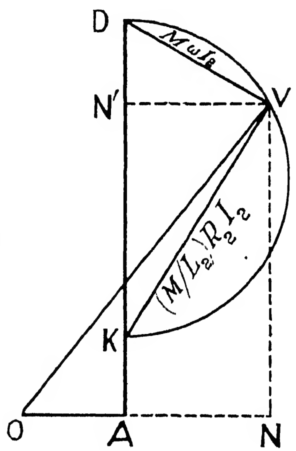


FIG. 191.

$$R' = R_1 + R_2\gamma^2,$$

where γ has been put for the ratio of secondary to primary current.

57. The effective reactance between the primary terminals is diminished by the presence of the secondary to the extent $N'D/I_1$. The similar triangles DN′V and DV′K give the proportions

$$DN' : N'V = DV : VK = L_2\omega : R_2.$$

Also the value of N′V is known as $\gamma^2 R_2 I_1$ from above. Therefore

$$N'D = \gamma^2 L_2 \omega I_1$$

Thus
$$L'\omega = L_1\omega - \gamma^2 L_2\omega$$
or
$$L' = L_1 - \gamma^2 L_2.$$

The ratio γ of secondary current to primary is an important quantity in high frequency transformers. It is easily expressed by a formula which may be deduced from the diagrams by expressing KD in terms of I_1 and also I_2. Or, directly, we may remark that the amplitude of the induced emf in the secondary is $M\omega I_1$ and that the impedance of the secondary is Z_2, and that therefore the secondary current is given by $I_2 = M\omega I_1/Z_2$, therefore
$$\gamma = I_2/I_1 = M\omega/Z_2$$
$$= M\omega/(L_2{}^2\omega^2 + R_2{}^2)^{\frac{1}{2}}.$$

The equivalent impedance of a transformer has been deduced by the analytical method in § 45.

Condenser Load.

58. We must now extend the above results so as to comprise a secondary circuit including a condenser, that is to say, a condenser load. Instead of Figs. 182 and 183 we shall have 192, 193 and 194. The reasoning is the same as before, but the secondary vector triangle DFB of Fig. 181 becomes the rhombus DPQB of Fig. 193. The vector marked $X_2 I_2$ is the excess of $L_2\omega I_2$ drawn in the direction of the arrow over $I_2/C_2\omega$ which is drawn in the opposite direction. In symbols:

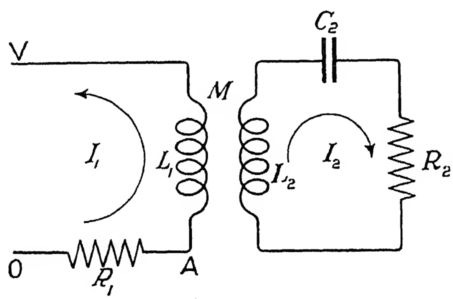


FIG. 192.

$$X_2 = L_2\omega - 1/C_2\omega.$$

Thus X_2 in the new diagrams will replace $L_2\omega$ in the old diagrams. This numerical difference does not alter the geometry of the diagrams.

Variation of Ohmic Load.

59. We may therefore pass at once to the circle of Fig. 194 for studying the variation of primary voltage with change of secondary resistance. In this Figure the assumption is that C_2 is of such a size that $1/C_2\omega$ is about the same magnitude as, though smaller than, $L_2\omega$ and thus X_2 is much smaller than $L_2\omega$. The consequence is that DK′ is greater than before and therefore the semicircle is greater. Thus with the same value of M and ω as before, DV is greater, which shows that I_2 is greater. Also, the image resistance of the secondary in the primary, namely, NA/I_1, is greater. It is also seen from the diagram that E makes a smaller angle with I_1, that is to say, the power factor is greater. It is to be noted that if the total reactance is negative the secondary voltage lies on the left of the diameter K′D. If it is zero F is at D for all values of R_2.

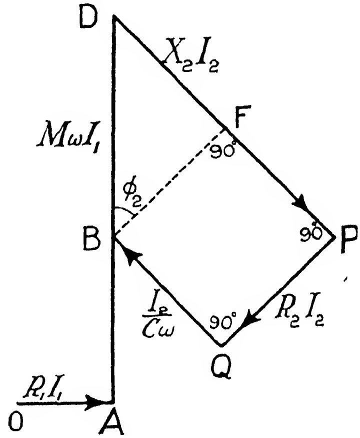


FIG. 193.

60. All the formulæ deduced in § 56 are valid and those of § 57 can be modified to suit the new circumstances by replacing $L_2\omega$ by X_2. In consequence the apparent resistance and inductance are given by

$$R' = R_1 + R_2\gamma^2$$

and

$$L' = L_1 - X_2\gamma^2\omega$$

where now

$$\gamma = I_2/I_1 = M\omega/Z_2 = M\omega/(X_2^2 + R_2^2)^{\frac{1}{2}}.$$

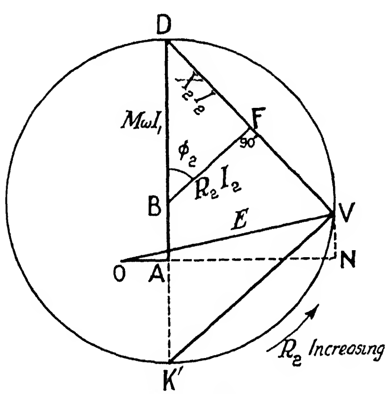


FIG. 194.

The numerical example of § 51 should now be reworked by the reader with a few values of C_2.

Variation of Reactive Load.

61. Let us now consider the problem, very important in Wireless Telegraphy, of the changes in currents and voltages brought about by variation of the capacitance from large values down to resonance and beyond. We have to imagine the circuit of Fig. 192 supplied with a constant primary current I_1 and the values of L_1, L_2, M, R_2 kept fixed while C_2 is varied from infinity (short-circuited condenser) to zero (open circuit). The reactance X_2 of course diminishes steadily during this process, passing through zero at resonance to negative values that gradually tend to infinity. The change of I_2 as C_2 varies and the consequent change of the reaction on the primary are of practical importance. In order to trace these variations we proceed as in § 21. Note first that in Fig. 195, because BF is always perpendicular to FD, the point F describes a circle on diameter BD as C_2 varies. But the vector $M\omega I_2$ is proportional to $R_2 I_2$ because M and R_2 are both constant, and it lies along DF and is therefore perpendicular to BF, and thus we conclude that as C_2 varies the point V distant $M\omega I_2$ from D describes a circle with diameter perpendicular to DB. This circle is shown in Fig. 195, which should be contrasted with Fig. 154 for direct coupling. As in that case, we determine the diameter of the circle locus from the consideration that ϕ_2 is zero when X_2 is zero, that is when resonance rules; on which occasion the secondary current I_2 is determined jointly by R_2 and the induced voltage $M\omega I_1$, that is to say

$$I_2 = M\omega I_1/R_2.$$

Therefore the diametrical vector DH is given by

$$\mathrm{DH} = M\omega I_2 = M^2\omega^2 I_1/R_2.$$

At this resonance adjustment the secondary current has of course its maximum value and has its vector perpendicular to BD; that is, is lagging a right angle behind the primary current.

If the capacitance C_2 be further reduced the remainder of the circle is described by V till when C_2 is zero the point D is reached. This means that the secondary is open and no secondary current is running. While V is in the upper semicircle the value of E attains its maximum, or, in other words, the effective impedance of the assemblage is a maximum. The exact adjustment is found graphically by joining O to the centre of the

circle and prolonging to cut the circle. The same line gives the

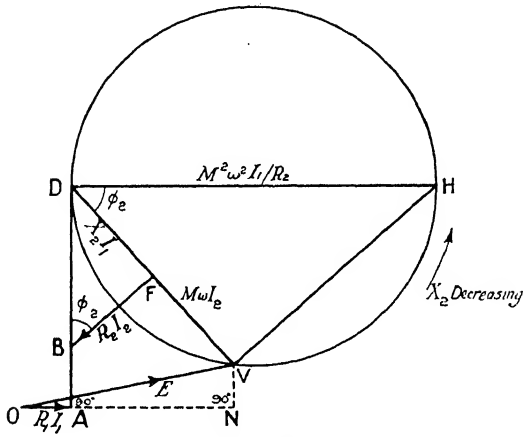


FIG. 195.

adjustment for minimum E, or the minimum effective impedance as shown in Fig. 197.

62. The semicircle of variable resistance with capacitance constant, such as in Fig. 194, must be carefully distinguished from the circle of variable capacitance with resistance constant, such as in Fig. 195. To emphasise the distinction, Fig. 196 is drawn showing both. The ends of both diameters, namely, H and K′, lie on a line through V perpendicular to the common chord DV. To each setting of C_2 a dif-

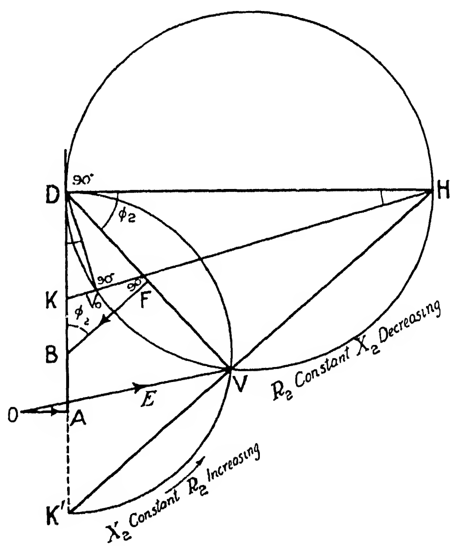


FIG. 196.

ferent resistance semicircle corresponds, and for each setting of R_2 there is a different capacitance circle.

To find the point on the capacitance circle representing the

short-circuited condenser, notice first that the angle VHD is equal to ϕ_2, whose tangent is R_2/X_2; and, second, that when C_2 is infinite $X_2 = L_2\omega$. The angle KHD has been made so that its tangent is $R_2/L_2\omega$ and the line HK therefore cuts the circle in the point V_0 required. Alternatively DK may be set off along DA of length k^2.DA, where k is the coefficient of coupling between the coils; then by joining K to H the point V_0 is determined.

Generalised Coupling Coefficient

63. The diagram of Fig. 196 brings out very plainly the fact that the diameter of the resistance semicircle may be greater than AD when there is condenser load. In fact the condenser may cancel as much as desired of the secondary inductance and thus make the effective coupling coefficient take any desired value. The diagram proves that DK′/DA $= M^2\omega/L_1X_2$, the square of the generalised coupling coefficient. Analogously to § 27 this coefficient may be called the coupling between the *circuits* in contradistinction to the coupling between the windings taken by themselves. The position of K′ is, in a numerical example, easily calculated from the formula just given.

Equivalent Resistance and Reactance.

64. The equivalent resistance and reactance between the primary terminals are given as in Figs. 195 and 197 by dropping a perpendicular VN on OA. The formulæ appear in § 60. Clearly ON is greatest, and therefore the equivalent resistance and the power absorbed likewise greatest, when V is adjusted to H, that is at resonance. All these are zero, on the other hand, when V is at D, that is on open circuit. The equivalent inductance has a minimum and a maximum when V is at the points of the circle

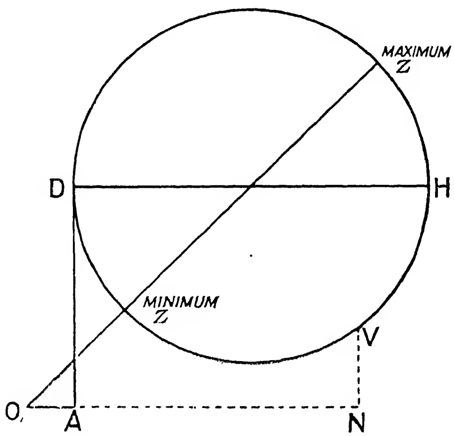


FIG. 197.

furthest from the diameter DH. In Fig. 197 the adjustments corresponding to the maximum and the minimum equivalent impedance also are found by joining O to the centre of the circle. At the place of minimum Z the transformer becomes in a sense "tuned" to the applied emf.

Some Special Adjustments.

65. Again, Figs. 195 and 196 point out that if R_2 is small enough the capacitance circle may be so large as to touch or to cut the prolongation of OA. This means that for a certain value or a certain pair of values of C_2 the vector OV may lie along OA, that is to say, the applied primary voltage may be in phase with the primary current, and the equivalent reactance between the

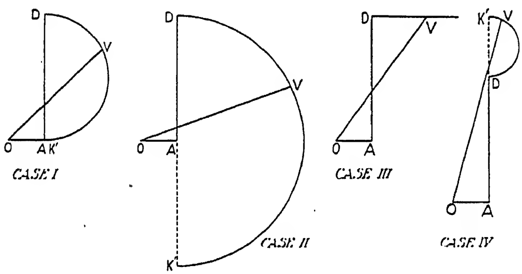


FIG. 198.

primary terminals will be zero. This has been treated fairly fully in § 29, and the reader may be left to adapt that discussion to indirect coupling.

The two circles of Fig. 196 comprise all the sine theory of the simple magnetic coupled circuit, and a great many interesting and practical problems could be studied by their aid if space allowed. It must suffice here to mention very briefly a few important cases arising out of the changes in size of the resistance semicircle. Notice first that if d be the ratio of the diameter of the resistance semicircle to DA, then when $C_2 = \infty$, $d = k^2 = M^2/L_1L_2$, while when C_2 has a finite value $d = M^2\omega/L_1X_2$. This leads immediately to the result

$$1/C_2 = \omega^2 L_2(1 - k^2/d)$$

from which the value of C_2 needed to make the resistance circle any desired diameter d.DA can be computed.

In case I, $d = 1$ and $1/C_2 = \omega^2 L_2(1 - k^2)$;
,, II, $d = 2$,, $1/C_2 = \omega^2 L_2(1 - \frac{1}{2}k^2)$;
,, III, $d = \infty$,, $1/C_2 = \omega^2 L_2$;
,, IV, $d = -k^2/a$ and $1/C_2 = \omega^2 L_2(1 + a^2)$.

These cases are illustrated in Fig. 198.

Case III is that of resonance. The most remarkable case is II, especially if the primary resistance is negligible; for in this event the end O of the primary voltage vector coincides with the centre of the semicircle, which implies that whatever the resistance in the secondary may be, from zero to infinity, the amplitudes of the primary voltage and current remain unchanged. The power factor alters however in response to the change of load.

Condenser in Primary Circuit.

66. A further extension of the circle diagram can at once be made to deal with problems in which there is a condenser in

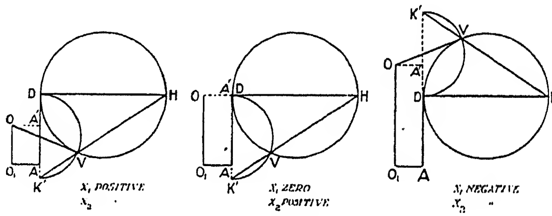


FIG. 199.

series with the primary winding. A series condenser merely annuls a portion of the whole primary inductance so that instead

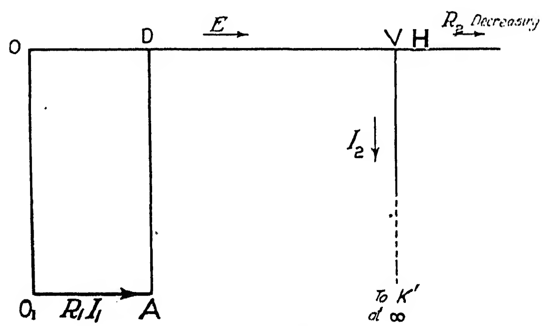


FIG. 200.

of the reactance $L_1\omega$ we have $X_1 = L_1\omega - 1/C_1\omega$, which may be positive, zero, or negative. Typical diagrams for the three cases

are given in Fig. 199. In all such diagrams the ratio DK′/DA′ is equal to $M^2\omega^2/X_1X_2$, which may be of either sign. A specially important instance is that of both circuits being "tuned" to the frequency of the alternator, when the diagram takes the simple form in Fig. 200. The resistance semicircle becomes the straight line DV, and V coincides with H. As R_2 decreases, the length DV increases. The applied voltage OV is in phase with the primary current and 90° ahead of the secondary current.

Mixed Self and Mutual Coupling.

67. In the transformers discussed up to § 32 the only coupling between the primary and secondary circuits is the direct coupling due to the common inductance L. But if an appreciable portion of the flux produced in any coil of either circuit threads a coil of the other circuit we obtain mixed inductive coupling. A case of some importance arises when the flux in the common coil effects linkages with another coil in the secondary circuit as indicated in Fig. 201.

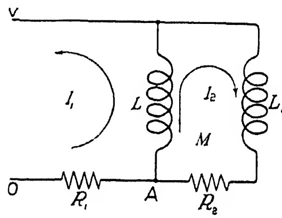


FIG. 201.

Let the mutual inductance between L and L_b be M. In order to draw the crank diagram (Fig. 202) we start with OA equal to R_1I_1, and draw AD equal to $L\omega I_1$. The latter vector is not the whole of the emf introduced by the primary current into the secondary circuit, for the contribution $M\omega I_1$ arises through mutual inductance. Set this off at D_1A. It is here drawn in a direction indicating that it assists the emf represented by AD, but whether it helps or hinders in fact depends on the relative positions of the coils and on the directions of the windings—briefly, M may be either positive or negative.

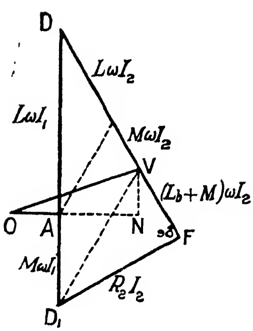


FIG. 202.

The total emf introduced into the secondary is therefore represented by D_1D and forms the hypotenuse of the secondary voltage triangle, the other two sides being DF and FD_1, equalling

respectively $L_2\omega I_2$ and $R_2 I_2$, These vectors are placed by making the angle at $D_1 = \angle L_2\omega/R_2$. It should be carefully noticed that the total inductance L_2 of the secondary circuit is made up of L_b from the load coil, L from the common coil, M linkages produced through L by current in L_b, with M linkages produced through L_b by current in L—a total of amount $L + L_b + 2M$.

Having obtained the vector DF, we now mark along it a length to represent the reaction of the secondary on the primary. This is made up of the effect of the indirect and the direct coupling. The length DV must in fact be equal to $(M + L)\omega I_2$, which is calculated by aid of D_1F already determined as $R_2 I_2$. The remainder VF of the secondary inductive vector will, of course, be $(L_b + M)\omega I_2$. The crank OV is the geometrical sum of the primary ohmic drop, the primary inductive drop, and the reaction from the secondary. This must be supported by the applied emf and therefore OV represents E, the emf needed in the primary to drive I_1 through the assemblage.

The equivalent resistance and inductance of the assemblage are obtained from ON and NV in precisely the same way as in § 56.

Inductance of Two Coils in Parallel.

68. The calculation of the resultant inductance of two coils possessing mutual inductance connected in parallel is a particular application of the above. As a rule the resistance of both coils may be neglected. We then obtain the one line diagram of Fig. 203. The vector NV may be represented as $L'\omega I_1$, where L' is the equivalent inductance. Thus

$$L'\omega I_1 = L\omega I_1 - (L + M)\omega I_2$$

or

$$L' = L - (L + M)(I_2/I_1).$$

In order to determine the current ratio I_2/I_1 we write D_1D in its two modes

$$D_1D = (L + M)\omega I_1 = (L + L_b + 2M)\omega I_2$$

or

$$I_2/I_1 = (L + M)/(L + L_b + 2M).$$

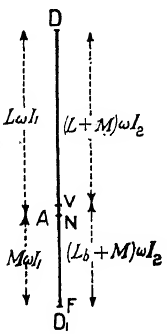


FIG. 203.

From this we deduce

$$L' = (LL_b - M^2)/(L + L_b + 2M),$$

in which M may be positive or negative.

Indirect Electric Coupling

69. Indirect coupling in various forms is illustrated in Figs. 204—206. The fundamental difference between indirect and direct coupling is the existence of conducting connection between the primary and secondary circuits in the direct class. Indirect coupling between two circuits is often unintentionally effected by the lines of electric force leaking from one condenser to the other.

Electric coupling has not had much application up to this

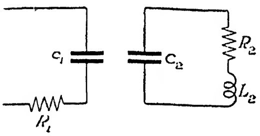


Fig. 204.

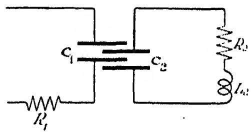


Fig. 205.

time, so it will be treated here much more briefly than was magnetic coupling. There is, however, a probability that electric coupling may find numerous applications in future C.W. work.

Turning to the problem presented by Fig. 204, it is clear that when the condenser c_1 is charged some of the Faraday lines of the charges cross from one plate of the condenser to the other by way of the plates of the condenser c_2, and thus the presence of c_2 will augment the capacitance of c_1, and reciprocally. These augmented capacitances will be given the symbols C_1 and C_2. In complicated circuits the rule previously given for estimating such augmented values should be used, namely, in the

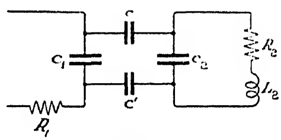


Fig. 206.

case of capacitance, the total capacitance C_1 is that active in the primary when the secondary is cut outside the coupling, and similarly for the capacitance C_2 of the secondary.

70. By writing down the equations of the circuits in full, it can be proved that a current of amplitude I_1 in the primary produces an emf in the secondary of amplitude $I_1/K\omega$ and that a current I_2 in the secondary produces an emf of reaction in the primary of amplitude $I_2/K\omega$. The quantity K may be called

the mutual capacitance. It is calculated from the self-capacitances and the augmented capacitances by the formula

$$\frac{1}{K^2} = \left(\frac{1}{c_1} - \frac{1}{C_1}\right)\left(\frac{1}{c_2} - \frac{1}{C_2}\right).$$

The vector diagram of primary and secondary is built up from the secondary and primary triangles by a step-by-step process analogous to those of §§ 33, 43. By combining the triangles we arrive at Fig. 207, where the dimensions of all the vectors are stated. The coefficient of coupling of the circuits is,

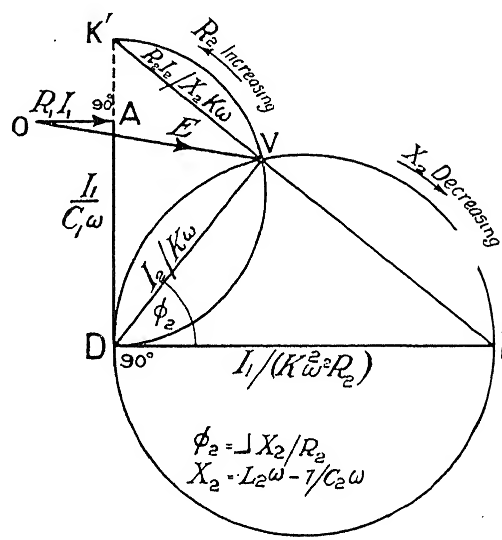


FIG. 207.

as before, given by the ratio DK'/DA, of which the value is $-C_1\omega/X_2K^2$. The meaning of this is most clearly seen if we take the particular case in which L_2 equals zero, for then $-X_2 = 1/C_2\omega - L_2\omega = 1/C_2\omega$ and the ratio is C_1C_2/K^2. This may be called the square of the coefficient of capacitance coupling.

Direct electric coupling may be regarded as a limiting case of the indirect, just as occurs in magnetic coupling. We need only imagine the separate condensers exactly equal and capable of superposition, and then suppose them moved into coincidence.

IRON CORE COILS AND TRANSFORMERS.

71. Within recent years alloyed iron has become commercially available in sheets of thickness less than 0·05 mm, and in consequence iron cored coils are beginning to appear in radio circuits. In suitably thin laminæ the losses due to eddy currents are relatively small and the hysteresis losses are negligible at low flux densities, and therefore it is possible to design transformers with iron cores in which the total losses are smaller than in air core transformers doing the same work; for the use of iron reduces the volume of copper employed and, consequently, the eddy current losses in the substance of the copper.

72. The permeability of iron even at low flux densities is at all frequencies more than 100 times that of air, and may be much greater than this at higher densities. But in order that the whole of the iron in a core shall be active it is necessary that it shall be finely divided; otherwise the skin-effect—by which is meant the shielding of the interior by induced currents flowing in the outer layers—puts much of the iron out of action and therefore lowers the average permeability. Assuming, however, that the core is all really active we see immediately that an iron core has two outstanding qualities; first, it increases the flux produced by unit current in a single turn of a winding, and second, it tends to lead a larger proportion of the magnetic lines from each turn through other turns and thus increases the linkages. Both qualities augment the self-inductance of a winding on the core and the mutual inductance between two windings on the same core. In particular, when the flux density is of such a value that the permeability takes values higher than 1,000 the leakage between turn and turn of a winding on a long cylindrical iron core is so small relative to the total flux that even a loose single layer winding, or a winding several layers deep, may be dealt with arithmetically as if it were a part of an ideal single layer solenoidal winding.

Skin Effect and Eddy Currents in Cores.

73. It is of advantage at this stage to look more closely at these phenomena. In Fig. 208 is seen a cross section of a slab of metal surrounded by a winding in which an alternating current is supposed to be flowing. An alternating magnetic flux perpendicular to the plane of the paper is therefore produced in the

metal, and the lines move sideways in and out of the metal during a half period of the alternation, being of opposite signs in consecutive half periods. As the flux cuts through the metal it induces, in accordance with the second law of electromagnetism, currents called Foucault or eddy currents, which flow in layers arranged like the coats of an onion. In fact the lump of metal acts as a secondary circuit to the coil. Evidently during a complete alternation an outer layer must be cut by more flux lines than an inner layer, and therefore the current induced in an outer layer is greater than that in an inner one. This would be true even if the flux were distributed uniformly throughout the metal, but in fact the flux is not so distributed. For since the outermost layers of current are in almost opposite phase to the current in the winding (being "secondary" induced currents) they tend to produce an opposing flux in the central portions of the metal

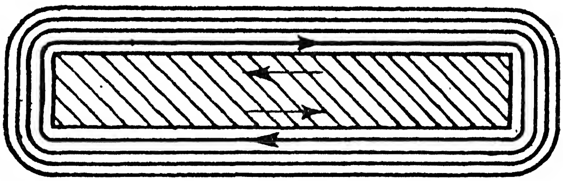


FIG. 208.

and therefore partially neutralise the flux that might have been produced by the winding. The magnitudes of the eddy current and of the flux density therefore fall rapidly from surface to centre, and both flux and current appear to be crowded into the outer layers of metal. Another consequence is that the total magnetic flux produced in a specimen of metal by a given magnetomotive force is smaller than it would be in the absence of the skin effect and therefore the apparent permeability is reduced.

74. For a large thin plate of metal the mathematical analysis is simple but is a little beyond the scope of this book. Not only does the amplitude of successive layers of eddy current vary from layer to layer but the phase changes gradually also. This is indicated by the crank diagram of Fig. 209, where each crank represents the current density in its plane. The cranks are supposed to be fixed to an axis OO′ which passes perpendicularly through the metal plate from one face at O to the other at O′, and they rotate with the axis at the angular velocity of the

alternating current in the winding. The instantaneous value of the current in any layer is obtained by projecting the length of its crank on a line in the layer and parallel to the reference line CA. A similar set of rigidly connected cranks can be drawn to represent the flux density in each layer. From the analysis of O. Heaviside and J. J. Thomson it may be concluded that one consequence of this distribution is that the P.D. at the terminals of the winding leads the current in the winding by less than $\frac{1}{2}\pi$ radians, and that at very high frequency the lead falls almost to $\frac{1}{4}\pi$. The exact value of the angle depends upon the thickness, the resistivity and the permeability of the metal.

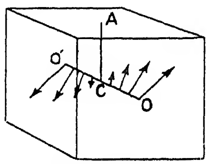


FIG. 209.

Application to Laminated Iron Cores. Eddy Losses in Thin Sheets.

75. These calculations should not be applied to iron without great caution. In iron the permeability varies with flux density owing to the non-linearity of the *B-H* curve, and also takes different values at the same value of *B* owing to hysteresis. Fortunately it is possible to make practical estimates of the losses due to eddy currents by combining simple reasoning with the results of experiments.

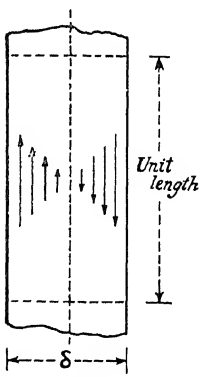


FIG. 210.

First let us suppose that in desiningg a coil for a given frequency the laminations of the iron of the core have been taken so thin that the flux density is nearly as great in the layers near the centre as at the surface, and let us consider how even finer lamination will affect the eddy current losses. Let Fig. 210 represent the section of a portion of a plate of thickness δ with eddy current flowing at the instant in the directions of the arrows. The resistance offered to the current will be proportional to the resistivity ρ of the metal and inversely proportional to the width of the path it flows in. Hence we have

$$\text{resistance} \propto \rho/\delta.$$

Again, the emf induced in the outer layer and producing the current in that layer is proportional to the amplitude of the flux entering and leaving the plate and to the frequency. A similar

statement holds regarding any layer. Hence, since the flux per unit length is equal to the product of the flux density B and the thickness δ we have

$$\text{eddy emf} \propto fB\delta.$$

By Joule's Law the rate of dissipation of energy is

$$(fB\delta)^2 \div (\rho/\delta)$$

or

$$f^2B^2\delta^3/\rho.$$

But the number of plates piled together to make a total thickness of 1 cm is inversely as δ, and therefore the eddy current energy loss per second per unit volume is given by

$$P_e \propto (fB\delta)^2/\rho.$$

The numerical factor lacking is to be supplied by experiment; a value is given for silicon iron (stalloy) in § 105.

The formula shows that the eddy current loss in an iron core of given dimensions is smaller the thinner the sheets; and in order to keep the losses low the thickness of the laminations must be taken smaller the higher the frequency of the alternations.

Eddy Currents in Round Wires.

76. Sometimes round wires are used rather than plates in the construction of cores. By reasoning similar to the above we may deduce an analogous formula for cases when the wire is so fine that the flux penetrates to the centre almost undiminished. For the length of the dotted path representing the mean eddy current in Fig. 211 is proportional to the radius r of the cylinder, and the area across which the current flows is also proportional to r. Therefore the resistance is independent of the radius and depends only on ρ. Again, the flux threading corresponding dotted circles in wires of different diameter is proportional to Br^2 and therefore

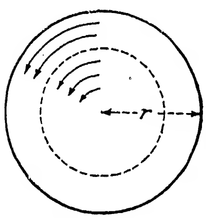


FIG. 211.

$$\text{eddy emf} \propto fBr^2.$$

Hence the rate of loss per cylinder is proportional to

$$(fBr^2)^2/\rho.$$

But the number of equal cylinders in a unit area of the cross section of a core is inversely proportional to the area of one cylinder, and therefore the rate of loss per unit volume is

$$P_e \propto (fBr)^2/\rho.$$

Thin and Thick Sheets.

77. The above formulæ are not to be applied when the skin effect is prominent, that is to say when the thickness of a lamina is so great, having regard to the frequency, that an appreciable proportion of the area of the cross section is put out of action by the shielding effect of the outer eddy currents. The state of things in this event is represented by the diagram of Fig. 212, where the small triangular areas associated with the surfaces of the plate represent by their ordinates the tapering values of the flux density in the skin. The depth of the skin, represented by the symbol y, will be defined as the distance from the surface at which the flux density becomes an assigned fraction, say 5 per cent., of its value at the surface—the flux beyond that depth will be neglected.

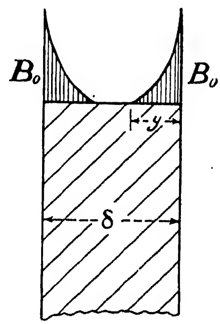


FIG. 212.

78. In order to obtain the complete formula for the thickness of the skin it would be necessary to use more advanced mathematics than is here desirable. But it is easy to obtain the most informative portion of the expression. For it is clear that the depth y may depend upon the frequency f of the alternations, upon the permeability μ of the metal, and upon the resistivity ρ of the metal; and that it cannot depend upon any of our other variables such as δ or B—for δ might be made infinite without affecting the skin and B is not concerned when the definition of skin thickness involves a ratio of reduction and not reduction to a specified absolute value. Therefore assume

$$y \propto f^a \mu^b \rho^c$$

where the indices a, b and c are unknown and must be determined. They may be determined by ensuring that the answer is in centimetres when in the assumed formula we substitute for f the number of alternations per second, for μ the value of the permeability in henrys per centimetre, for ρ the value of the resistivity in ohm-centimetres, as directed by II, § 112. We do not need for this purpose to use any particular numerical values, and so, ignoring numbers, we write

$$\text{cm} = (\text{sec})^{-a}\,(\text{henry/cm})^b\,(\text{ohm-cm})^c.$$

But $\qquad$ henry = sec-ohm.

Therefore $\qquad \text{cm} = (\text{sec})^{-a+b}(\text{cm})^{-b+c}(\text{ohm})^{b+c}.$

This can be satisfied only by making

$$-a+b=0,\ -b+c=1,\ b+c=0,$$

from which we obtain

$$a=-\tfrac{1}{2},\ b=-\tfrac{1}{2},\ c=\tfrac{1}{2}.$$

Hence finally $\qquad y \propto \sqrt{\frac{\rho}{\mu f}}.$

This shows that the penetration is greater the greater the resistivity, the smaller the permeability, and the smaller the frequency.

79. The missing numerical factor can be obtained only by more elaborate analysis; it is 0·477 approximately. Hence

$$y = 0{\cdot}477\sqrt{\frac{\rho}{\mu f}}$$

gives the depth at which the eddy current density and flux density become reduced to 5 per cent. of their surface value. Since the depth is given by this formula in centimetres, the formula is true for em units and for practical units. It must be mentioned that in the formula the permeability μ is assumed to be a constant, which is not the case in iron when a certain value of flux density is exceeded. Leaving this aside for a moment the formula enables us to decide whether to call a sheet "thick" or "thin"; it is thick when its thickness exceeds $\sqrt{(\rho/\mu f)}$ and we call a sheet thin when its thickness is decidedly less than this. The remark on the variability of permeability shows that a plate regarded as thin at low flux densities might have to be regarded as thick at high flux densities on account of the reduced depth of the skin at the higher permeability. This can be seen by evaluating the quantity $\sqrt{(\rho/\mu f)}$ at, say, 10,000 $\sim$, for $\mu = 1{,}000$ and for $\mu = 150$, taking in both cases $\rho = 10{,}000$ emu. With the larger value of the permeability we obtain $y = 0{\cdot}010$, with the smaller value $y = 0{\cdot}026$.

Eddy Current Losses in Thick Sheets.

80. Let B' represent the average or apparent flux density in the case shown in Fig. 212 and B_0 the density at the surface, then

$$B'\delta = B_0 y$$

if we assume for the purposes of an easy approximation that the flux density tapers in a linear manner. We have also

$$\text{eddy emf} \propto f B_0 y$$
$$\text{resistance} \propto \rho/y,$$

and therefore the rate of loss per square cm of plate is proportional to

$$(fB_0y)^2y/\rho,$$

and the rate of loss per cubic cm of metal is

$$P_e \propto (fB_0y)^2y/\rho\delta$$

or

$$(fB'\delta)^2y/\rho\delta.$$

If the apparent flux density B' is specified we use the latter result, if the magnetomotive force is given we use the former. The condition B' constant corresponds to constancy of the terminal voltage of the winding; while B_0 constant corresponds to the constancy of the ampere turns. In either case we substitute for y and obtain

$$P_e \propto \frac{B_0^2}{\delta}\sqrt{\frac{f\rho}{\mu^3}}$$

or

$$P_e \propto B'^2\delta\sqrt{\frac{f^3}{\mu\rho}}.$$

81. Thus if a particular coil is supplied in turn with cores built up of sheets of different thicknesses of the same iron, the eddy losses will be greater with the thicker laminations if the coil be energised at constant terminal voltage, but will be smaller with the thicker laminations if the current through the coil be of constant amplitude. In the constant voltage case the losses increase much more rapidly with increase of frequency than in the constant current case. These results, however, do not appear to have been confirmed as yet by experiments on laminated cores at high frequencies.

82. The above formulæ hold good for the losses in cores built up of round wires if the radius of the wire be used in place of the thickness δ of the sheets.

Hysteresis Loss.

83. When an alternating current flows in a coil with iron core the flux density in any particular portion of the metal attains in turn extreme values of equal magnitude and opposite sign denoted by the symbol B in this paragraph. In a complete cycle of changes the flux density returns to its original value and a definite quantity of heat, so it was experimentally discovered, is developed in the iron. The heat generated is a function of the extreme value of the flux density attained in the cycle and is the same whether the cycle is performed slowly or quickly, so far as is known. This phenomenon is one aspect

of the larger phenomenon called hysteresis, which is discussed from another point of view later. Measurements of the heat developed show that, for example, when a sample of iron is taken steadily from a flux density of 10,000 lines per cm^2 to a density of $-10{,}000$ and back again, the energy value of the heat liberated is from 0·00015 to 0·00025 joule per cm^3 in silicon iron, and at least 50 per cent. greater in ordinary annealed transformer iron. From actual measurements of hysteresis loss Steinmetz obtained the empirical formula

$$\text{Loss per cycle per cubic cm} = \eta B^{1 \cdot 6}.$$

Hence η is called the hysteresis coefficient; its value ranges from 0·0006 to 0·001 for silicon iron and from 0·001 to 0·0022 for annealed transformer iron, when the electromagnetic system of units is used in the above equation. Plainly if f cycles are performed per second the rate of dissipation of energy will be given by

$$P_h = f\eta B^{1 \cdot 6}\, 10^{-7} \text{ watt per cm}^3$$

where B is to be expressed in " lines."

Application to Laminated Cores.

84. The hysteresis loss when the plates are thin in the sense defined in § 77, does not depend upon the absolute thickness of the laminæ, for in this case the flux penetrates easily throughout the whole substance of the iron; in other words, the fineness of division of the iron has no effect, and the equation last written therefore suffices.

85. When the sheets are so thick that the central layers are shielded effectively by the eddy currents we may write the rate of loss of energy per cubic centimetre in the first place in the form

$$P_h \propto f\eta B_0{}^{1 \cdot 6} y/\delta$$

because only the fraction y/δ is utilised (§ 77). This takes two forms according as we have to regard the surface value of the flux density as fixed or the average value as fixed. In the former case we have finally

$$P_h \propto \frac{\eta B_0{}^{1 \cdot 6}}{\delta}\sqrt{\frac{f\rho}{\mu}}.$$

When, on the other hand, the apparent density is fixed we have, since $B'\delta = B_0 y$,

$$P_h \propto \eta B'^{1 \cdot 6} f^{1 \cdot 3} (\delta^2\mu/\rho)^{0 \cdot 3}.$$

The formulæ for round wires are the same as for plane sheets.

No experimental results appear to have been published as yet that will enable the missing numerical constants to be determined in the last few formulæ.

Total Losses.

86. In any experiment upon core losses the eddy losses and the hysteresis losses are lumped together and the sum of P_e and P_h is observed. At high frequencies the plates are apt to be "thick" in the sense of § 77 and consequently the sum of P_e and P_h is a rather complicated function of δ, f, ρ, μ and the experimental conditions. Some divergence is therefore to be expected between the conclusions of different observers. The likelihood of this is increased when it is recalled that the permeability μ increases greatly when the mmf increases beyond a certain value.

For silicon iron in thin plates we have approximately

$$\left.\begin{array}{l}\text{Total losses}\\ \text{per cubic cm}\end{array}\right\} = 0{\cdot}9(fB'\delta)^2/\rho + 7 \times 10^{-4} fB'^{1{\cdot}6}$$

in ergs per second,

the values of B', ρ and δ being in em units.

For iron in thick sheets the formulæ are of the form

$$\alpha_0 \frac{B_0^2}{\delta}\sqrt{\frac{f\rho}{\mu^3}} + \beta_0 \frac{\eta B_0^{1{\cdot}6}}{\delta}\sqrt{\frac{f\rho}{\mu}}$$

when the surface value of the flux density is given, and

$$\alpha' B'^2 \delta \sqrt{\frac{f^3}{\mu\rho}} + \beta' \eta B'^{1{\cdot}6} f^{1{\cdot}3}\left(\frac{\mu\delta^2}{\rho}\right)^{0{\cdot}3}$$

when the apparent flux density is given. It is noticeable that when the surface value of the flux density is given the total losses may be expected to vary as

$$\frac{\sqrt{f\rho}}{\delta},$$

and are therefore reduced by finer lamination and increase only slowly with rise of frequency.

For stalloy the values of α_0 and β_0 have been approximately determined as $\alpha_0 = 0{\cdot}3, \beta_0 = 0{\cdot}15$, all units being in the em system.

Flux in a Solenoid.

87. The formula for the uniform magnetic field in an infinite solenoid is given in II, § 52 as

$$H = 4\pi(\tau/l)i$$

where H, the magnetic force, and i the current, are in emu and τ is the number of close equal turns in a length l cm of the solenoid. The flow of magnetic lines per square centimetre of cross section of the solenoid, in other words, the flux density B, is given by $B = \mu H$, where μ is the average permeability of the core, and the magnetic flux through the core is, if A be the area of cross section in sq. cm, $\mathcal{F} = AB$. It is usual to express the flux $\mathcal{F}$ in absolute units, even though the current be expressed in practical units, and so we have the equation

$$\mathcal{F} = \mu AH = 0{\cdot}4\pi\mu(\tau/l)Ai$$

where i is in amperes.

88. For engineering calculations a slightly different point of view is sometimes helpful. A current in a winding such as

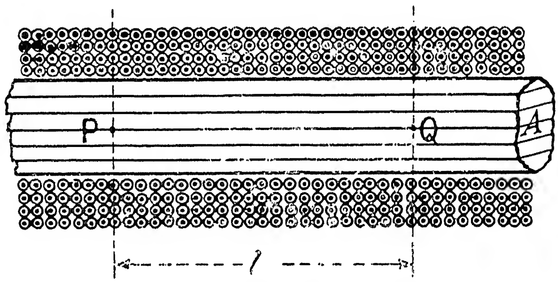


FIG. 213.

indicated in Fig. 213 is regarded as producing a magnetic pressure or magnetomotive force between any two places such as P, Q, this mmf being the line integral of the magnetic force from P to Q, namely, $H \times l$. Using the above formula for H we have

$$\begin{aligned}\text{mmf from P to Q} &= 0{\cdot}4\pi i\tau \\ &= 1{\cdot}257 \times \text{ampere turns between P and Q.}\end{aligned}$$

89. In a magnetic circuit magnetomotive force produces a flux of magnetic lines just as emf produces a flow of current in an electric circuit. The magnetic analogue of the resistance of an electric circuit is called reluctance, and just as in the electric circuit we compute the resistance of a part length l of uniform sectional area A by aid of the expression $l/\gamma A$ (where γ is the conductivity), so we compute reluctance by means of the expression $l/\mu A$ for each uniform portion. We finally arrive at the

equation for the flux $\mathscr{F}$ between P and Q of Fig. 213 in the form

$$\text{flux} = \text{mmf} \div \text{reluctance}$$

or

$$\mathscr{F} = 0{\cdot}4\pi i\tau \div l/\mu A$$
$$= 0{\cdot}4\pi\mu(\tau/l)Ai.$$

90. For a complete magnetic circuit in which there are cores, yokes, joints and air gaps of varying sections and lengths the total reluctance is the sum of the separate reluctances. As an application of this principle we may write down the equation for the ampere turns needed to force a given flux $\mathscr{F}$ (in emu) through a magnetic circuit in the form

$$i\tau = 0{\cdot}8\ \mathscr{F}(l_1/\mu_1 A_1 + l_2/\mu_2 A_2 + \ldots\ldots)$$

where the suffix 1 refers to the first portion, suffix 2 to the second portion, and so on.

Inductance of Coils with Iron Cores.

91. As already remarked, the magnetic leakage between turns of a coil on a long or a closed iron core may be neglected if all the copper wire is reasonably near the iron. This makes the approximate calculation of self-inductance very simple—we calculate the linkages per emampere by multiplying the flux by the number of turns. For the portion PQ of Fig. 213 the flux in emu is $4\pi\mu(\tau/l)A$ per absolute unit of current (10 amperes) and therefore

$$L \text{ in emu} = 4\pi\mu(\tau/l)A \times \tau$$
$$= 4\pi\mu\tau^2 A/l.$$

Now 10^9 emu = 1 henry. Hence

$$L \text{ in practical units} = 1{\cdot}25 \times 10^{-8}\mu\tau^2 A/l.$$

This simple formula gives quite good results for, say, a choking coil. The symbol l refers to the length of the closed magnetic circuit because it entered the equation in the computation of the reluctance; it does not refer to the length of the winding, which may in fact cover only a portion of the length of the core. In a transformer there are two windings, primary and secondary, on the same core, and therefore we have

$$L_1 = 1{\cdot}25 \times 10^{-8}\mu\tau_1^2 A/l,$$
$$L_2 = 1{\cdot}25 \times 10^{-8}\mu\tau_2^2 A/l,$$

if the core is of uniform section throughout. If it is not we must take $l/\mu A$ to be the sum of the reluctances for the separate portions. In order to calculate the mutual inductance between primary and secondary we take the flux due to unit current in

either and multiply by the number of turns in the other winding. Taking the current to be in the primary we obtain

$$M \text{ in emu} = 1{\cdot}25\mu\tau_1(A/l) \times \tau_2$$

$$M \text{ in practical units} = 1{\cdot}25 \times 10^{-8}\mu\tau_1\tau_2 A/l.$$

92. From the definition of the coefficient of coupling in II, § 64, we obtain

$$k^2 = M^2/L_1L_2 = 1$$

and therefore $\sigma = 1 - k^2 = 0.$

That is to say, the assumption that the magnetic leakage is zero is bound up with these formulæ. As a matter of fact the value of k^2 in a well-designed power transformer may be as large as 0·95, and therefore the leakage coefficient 0·05. It is legitimate to use the above formula as an approximation for M whenever the subsequent mathematical operations do not bring it into subtraction from the product L_1L_2, in which case it is better to use

$$M = 1{\cdot}25 \times 10^{-8}\mu k^2\tau_1\tau_2 A/l.$$

It is not possible to calculate σ accurately direct from the geometrical data of the coil even when the permeability is supposed constant, but as σ itself usually appears as a small correction its exact value is not often demanded. For simple approximate methods of estimation books on transformer design should be referred to. But the value of σ depends greatly on the permeability, which in turn depends on the flux density; the leakage is much greater when the flux density is large or small than when it has that medium value at which the permeability is a maximum. The coefficient of magnetic leakage of a transformer varies, in fact, periodically during the cycles of an alternating current traversing its windings. The self and mutual inductance themselves vary similarly for the same reason; yet it is customary and proves sufficiently good for practical purposes to take these as constant, leaving the departure from constancy to be allowed for if at all by the harmonics introduced by such variations into the fundamental alternation.

Crank Diagrams.

93. By the aid of the approximate formulæ above, the standard transformer crank diagram may be viewed somewhat differently. In Fig. 214, which is the crank diagram for transformers with negligible primary resistance, the vector AD represents the voltage amplitude $L_1\omega I_1$ and we know that L_1I_1, the primary

flux linkages, is equal to $\tau_1 . \mathscr{F}_1$ where $\mathscr{F}_1$ is the flux in the core due to the primary current in its winding. Again VD is $M\omega I_2$, and we know that $MI_2 = \tau_1 . \mathscr{F}_2$, where $\mathscr{F}_2$ is the flux due to the secondary current in its winding. Now AV is the geometrical sum of AD and DV. It therefore represents $\omega\tau_1 . \mathscr{F}$, where $\mathscr{F}$ is the resultant flux in the core. In words AD, DV and AV are all voltages induced in the primary turns by component fluxes threading those turns, AD being due to the component flux produced by the primary current, DV to the component flux produced by secondary current, and AV to the flux existing in the core as the resultant of the former two and cutting the primary turns. But this vector is the primary applied voltage E (in the absence of primary resistance and other causes of loss), and therefore we conclude that the whole flux actually existing in the core under any condition of loading is always of such magnitude as to produce in the primary winding a counter emf that nearly balances the primary applied emf. It is necessary to say "nearly," because in general the copper and iron losses absorb a portion of the primary applied emf.

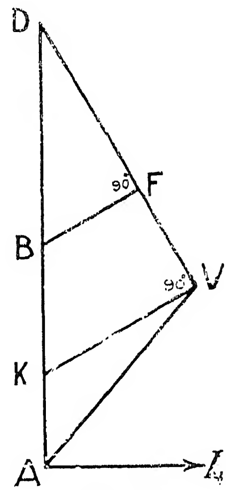


FIG. 214.

Having thus accounted for the counter emf in the primary winding, we turn to the secondary voltage FB, which is the same as the voltage R_2I_2 applied to the load, and inquire if it can be similarly derived from the linkages it makes with the core flux or a part of it. But for the sake of proper comparison with the primary, account should be taken of the fact that the primary has τ_1 turns and the secondary τ_2. If this be done by multiplication by τ_1/τ_2 we obtain instead of FB the parallel vector VK; as can be proved by noticing that $VD : FD = M : L_2 = \tau_1\tau_2 : \tau_2^2$. In fact KV measures the flux affecting the secondary on the same scale as AV measures that affecting the primary turns. If there were no magnetic leakage these fluxes would be equal in magnitude and phase, but not otherwise. The diagram gives the difference quantitatively. For AV is geometrically equal to AK and KV; that is to say AV, the rate of change of flux linking

with the primary, is equal to KV, the rate of change of the flux linking with the secondary, together with AK, the rate of change of the flux failing to link with the secondary, that is, the leakage flux. All these rates of change are to the same scale because the voltages are reduced to the primary for the sake of numerical comparison. The point K is determined, it will be remembered, by the equation AK/AD $= \sigma$, the coefficient of magnetic leakage. In order to obtain vectors to represent the main flux and its components in phase as well as magnitude it is necessary to rotate the triangle AKV clockwise through a right angle—in its present position this triangle represents the rates of change of the flux linkages.

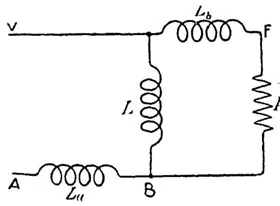


FIG. 215.

94. This division of the whole magnetic flux into portions is carried by some designers farther than in the last paragraph. The process, which is admittedly very artificial, may best be illustrated by considering a one-to-one transformer. Evidently not all the primary lines thread the secondary turns nor all the secondary lines the primary turns, but it may be assumed that a common portion of the two sets threads both circuits. Let the common portion be imagined to belong to a common coil of self-inductance L, the primary leakage lines to belong to a detached coil of inductance L_a, and the secondary leakage lines to belong to a detached coil L_b, as indicated in Fig. 215. The crank diagram for these circuits is sketched in Fig. 216. Here BD $= L\omega I_1$, DV $= L\omega I_2$, AB $= L_a\omega I_1$, VF $= L_b\omega I_2$, FB $= R_2 I_2$ and AV $= E$. All three coils are designed with the same number of turns.

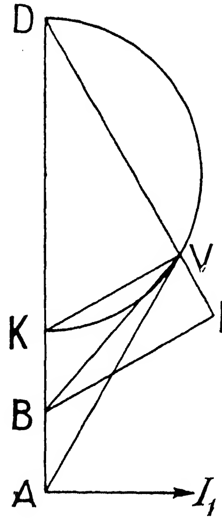


FIG. 216.

It is clear that the vector BD will represent the flux produced in the common coil by the primary current to the same scale as the vector DV represents the flux produced in the common coil by the secondary current; and therefore the resultant flux in

the common coil is represented to that scale by the resultant vector BV. Also AB is proportional to that part of the primary flux not linking with the secondary circuit while VF is proportional to that part of the secondary flux not linking with the primary circuit. We may therefore read the diagram of Fig. 216 as follows: The emf AV applied to the primary terminals is used in supplying the reactive potential drop AB and in overcoming the counter emf BV due to the common flux; the common flux, in its relation to the secondary circuit, generates the emf VB, which is used in supplying the reactive potential drop VF and the ohmic drop FB in the winding and the load.

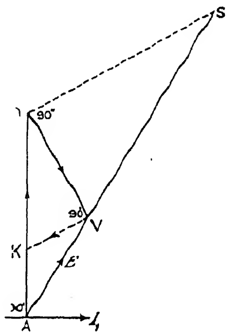


FIG. 217.

Equivalent One-to-one Transformer.

95. The general case of the transformer of any ratio can be reduced to an equivalent one-to-one transformer in several ways; a simple and useful instance arises when the load in the secondary circuit is purely ohmic and the resistance of the secondary winding negligible. In this case the secondary triangle BFD of Fig. 216 can be reduced by mere multiplication of its sides by a factor to the triangle KVD, and therefore the given transformer can be reduced to the one-to-one transformer of Fig. 215 with the secondary inductance L_b omitted. The constants of the primary circuit and its current and voltage will be unaffected, but the secondary resistance, current and voltage of the equivalent transformer will be different from those of the given. Let dashes indicate the quantities in the equivalent transformer, then we have from the triangle

$$V_2' = (M/L_2)\, V_2$$

or

$$R_2'I_2' = (M/L_2)\, R_2 I_2,$$

when M and L_2 relate to the given transformer. Let it be required that the consumption of energy in the secondary of the equivalent transformer shall be the same as that in the given; then

$$R_2' I_2'^2 = R_2 I_2^2$$

From these equations

$$I_2' = (L_2/M)I_2$$

and

$$R_2' = (M/L_2)R_2$$

We have also

$$L_a = \sigma L_1$$
$$L = k^2 L_1.$$

The equivalent transformer obeying these equations, which is really the shunt circuit of Fig. 215 when $L_b = 0$, represents the given transformer for all values of the ohmic load.

Applications of this theorem will appear later.

Constant Voltage Diagram.

96. The semicircle in Fig. 216 is drawn on the assumption that the primary current is kept constant, the applied primary voltage being varied to maintain this constancy. In power transformers it is the applied voltage that is kept approximately constant. Special circle diagrams for the elucidation of problems subject to this condition may easily be deduced from the above diagrams for the cases in which the primary resistance is negligible. Such a case is represented in Fig. 217. In this Figure AV is prolonged to S and DS is drawn parallel to KV. Then the triangles ADS and AKV are similar and therefore

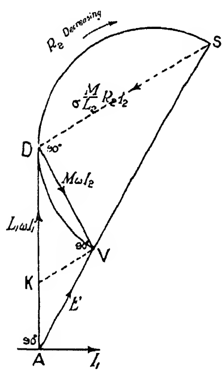


FIG. 218.

$$\frac{AD}{AK} = \frac{AS}{AV} = \frac{DS}{KV}$$

or

$$\frac{1}{\sigma} = \frac{AS}{E} = \frac{DS}{MR_2I_2/L_2}$$

or $AS = E/\sigma$ and $DS = MR_2I_2/L_2\sigma$.

These lengths are transferred to Fig. 218 in which if AV be kept constant AS and VS are also constant if σ is. Therefore, since the angle VDS is a right angle, variation of R_2 causes D to move along the new semicircle drawn in the Figure. When R_2 is great D is near V, when it is small D is near S. In fact the points V and S represent the open and the short-circuit conditions. If now every line in the Figure be divided by $L_1\omega$ and the whole diagram rotated backwards through a right angle (as is suggested by division by ω) we obtain Fig. 219. Here the primary current is represented by AD, its wattless component

(called the magnetising current I_μ of the transformer) by AV and the secondary current is proportional to DV. The length of AS is I_μ/σ. The smaller the magnetic leakage the greater the radius of the semicircle, and when the leakage is zero the semicircle becomes a straight line drawn from V perpendicularly to AV. The reader is referred to Russell's *Alternating Currents* for further information. It is sufficient for our present purposes to notice that if we require to trace the variations of, for instance, the value of the secondary terminal emf when the primary voltage is constant we need only take the *ratio* of BF to AV, or, subject to a factor, the *ratio* of KV to AV when using constant I_1 diagrams; and we must take the crank SD when using constant E diagrams. The percentage falling-off of the value of this voltage as the resistance changes from no load to full load is called the "regulation" of the transformer. The percentage is easily seen from the diagrams to be greater the greater the magnetic leakage.

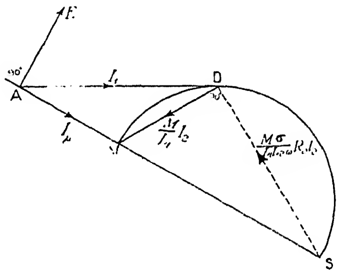


FIG. 219.

97. When the primary resistance or other loss is not negligible we use the crank diagram of Fig. 220. The primary applied voltage is OV $= E$ which may be parted into a loss component OA and the flux-balanced component voltage AV. The regulation (§ 95) has then to be computed by use of the ratio BF/OV. That is to say, the value of this ratio is found when F and V are at the no-load point D, and also when F and V are at the full-load point. The difference is expressed as a percentage of the former. The designer usually arranges that the full-load point comes nearer K than D on the semicircle. If the secondary copper resistance is not negligible it must be supposed included in R_2 and set off to scale at FF′; the regulation is then computed by use of the ratio BF′/OV. The angle FDF′ is constant at all non-reactive loads. Graphical methods are much easier than algebraical in such calculations.

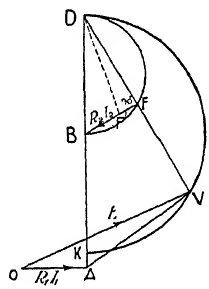


FIG. 220.

Emf Induced in Winding.

98. In the above discussion the emf induced in a winding by sine variation of a flux linking with the winding has often been referred to. It is of magnitude $L\omega I$ for a single coil of self-inductance L, where I is the amplitude of the current, as has been proved already. But this expression is used in another form by machine designers. Let $\mathscr{F}_m$ be the maximum value of the flux in the core of a solenoidal coil of τ turns, then by Faraday's Law $\quad e$ in volts $= -D(\tau . \mathscr{F}) \times 10^{-8}$
where $\mathscr{F}$ the flux at any instant is given by

$$\mathscr{F} = \mathscr{F}_m \sin \omega t.$$

Therefore

$$e = -\tau \mathscr{F}_m \omega \cos \omega t \, . \, 10^{-8}$$
$$E_m = 2\pi\tau f \mathscr{F}_m \, . \, 10^{-8}$$
$$E_{rms} = \sqrt{2}\pi\tau f \mathscr{F}_m \, . \, 10^{-8}$$
$$= 4 \cdot 44 \tau f \mathscr{F}_m \, . \, 10^{-8}.$$

Transformer Design

99. Only the merest outline need be given here as there are many excellent special books on the subject. All these are

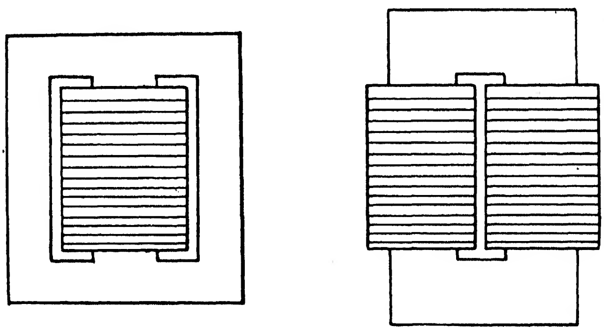


Fig. 221.

aimed, however, mainly at power transformer design, and the conditions studied are very different in detail from those met with in the new and unstandardised art of iron core transformer design for radio frequencies. In power plant there appear two types of transformer, core type in which the iron appears as the core of the copper winding, and shell type in which the iron is arranged as a shell round the copper winding. Simple forms of these are shown in Fig. 221. In one the copper is disposed round

the iron, in the other the iron round the copper. In our summary view of the principles of design we shall suppose for the purposes of calculation that the iron core is cut through and straightened out with its windings on it without losing the advantages of a closed magnetic circuit. In this way we evade the algebra that arises at the corners and obscures the issue by its tediousness.

The primary voltage, the voltage ratio, the frequency and the volt-ampere rating are, of course, given ; the second implies that the ratio of secondary to primary turns is given. Besides these the efficiency at various loads, and the regulation, may have limits assigned to them.

100. Let Fig. 222 represent the straight core with its primary and secondary copper windings. Let the sectional area of the

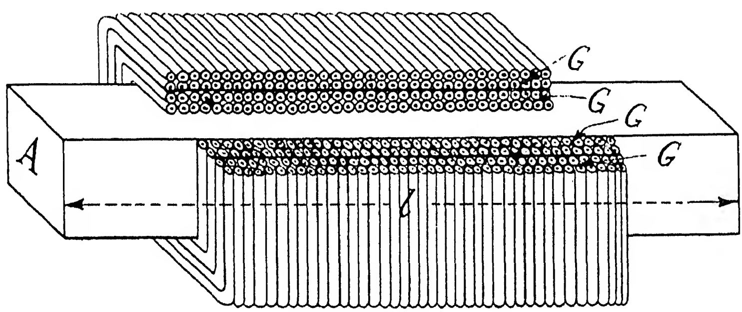


FIG. 222.

alloyed iron core with its insulation between the laminæ be A and the cross section of the copper and insulation of the primary winding be G. We shall suppose the leakage small and then AD $\doteqdot$ VD in Fig. 214. Since AD $= L_1\omega_1 I$ which is proportional to $\tau_1^2 I_1$, and VD $= M\omega I_2$, which is proportional to $\tau_1\tau_2 I_2$, we see that $\tau_1 I_1 \doteqdot \tau_2 I_2$; or, in words, the primary ampere turns approximately equal the secondary ampere turns under heavy load.

101. The designer can by aid of experience fix at the start the current density allowable in the copper, that is the current in rms amperes per sq. cm of cross section of the wire, account being taken of the method of cooling. Let this be Δ after allowing for the space factor introduced by the use of insulation, and suppose that Δ is the same for both primary and secondary wires. Then the primary ampere turns, being equal to the whole current

encircling the iron, is expressible as ΔG. Hence we obtain, in rms measure,

$$\tau_1 I_1 = \Delta G.$$

The secondary ampere turns being equal to the primary at full load, we conclude that the section of the secondary insulation and copper should be equal to that of the primary.

The designer also appeals to experience in fixing the limit permissible to B the flux destiny, after allowing for the space factor of the lamination. Since the losses in iron are usually recorded in terms of the maximum flux density attained in the cycle we take B to mean the maximum not the rms value. Then, by the definition of flux density, we have $\mathcal{N}_m = BA$. Hence from § 97, in rms measure, we have

$$E/\tau_1 = 4{\cdot}44 fBA \,.\, 10^{-8}.$$

This expresses the fact that the quantity called "the volts per turn" is proportional to the frequency and the core flux jointly.

102. The last two important equations are the essential electromagnetic equations for design. Sometimes it is convenient to multiply these equations to obtain

$$EI_1 = 4{\cdot}44 fB\Delta AG \,.\, 10^{-8}.$$

which states that the product of the sectional areas is a constant for a given volt-ampere rating and frequency. This is used in place of one of the pair. But there are three unknowns, namely, G the area of the cross section of each winding, A the area of the alloyed iron core, and τ_1 the number of primary turns. For definite solution of the equations a third equation is required.

103. There are many ways of supplying the necessary equation or condition. For instance it may be demanded that the losses shall be a minimum, or the initial cost of manufacture a minimum; and different rules of design have been standardised to meet the different requirements. An excellent discussion of many of these modes will be found in Denton's *Commercial Transformer Design.* Experience gained on completed transformers is of course an invaluable, not to say indispensable, guide in all the successful methods. Perhaps the easiest and clearest is that proposed by J. K. Catterson-Smith and discussed by him in *The Electrician* of January 3, 1913, of which the essence is to take from existing practice an approximate value of the ratio of the winding section to the core section. Let this assumed

ratio be represented by r. Then we have the following three equations to solve :—

$$\tau_1/G = \Delta/I_1$$
$$AG = \text{a known quantity, say } P$$
$$G/A = r$$

from which we obtain

$$G^2 = Pr$$

or $$G = \sqrt{Pr}$$
$$A = \sqrt{(P/r)}$$

and $$\tau_1 = (\Delta/I_1)\sqrt{Pr}.$$

After this stage minor geometrical considerations may be called upon to settle the linear dimensions of the sectional areas A and G, the more important being the obvious one that the window of the magnetic circuit must be of area large enough to accommodate the primary and secondary copper—a fact which partially fixes the length of the magnetic circuit. Another consideration arises from the circumstance that the outer layers of the windings should not be disproportionately far from the core, which tends to make the coils long rather than deep. After these details are decided the iron losses are easily calculated by aid of empirical expressions embodying the results of tests on hysteresis and eddy current phenomena, and the copper losses by aid of Joule's equation. An example will make this clear.

Numerical Example.

104. Let a preliminary design be required for a core type transformer of 20 kilovolt-ampere capacity for raising voltage from 1,000 to 10,000 at a frequency of 1,000 $\sim$. Let the current density be taken as 300 amperes/cm² and the space factor as 1/3 for both windings; so that $\Delta = 300 \times 1/3 = 100$ in rms measure. The copper is in the form of finely stranded cable. Let the maximum flux density be 8,000 lines/cm² and the space factor be 0·80, so that $B = 8{,}000 \times 0{\cdot}80 = 6{,}400$ lines/cm². The volt-ampere equation, § 101, gives

$$\begin{aligned} AG &= EI_1 \,.\, 10^8 \div 4{\cdot}44 fB\Delta \\ &= 20 \times 10^{11} \div (4{\cdot}44 \times 1{,}000 \times 6{,}400 \times 100) \\ &= 703{\cdot}8 \end{aligned}$$

Again, since $EI_1 = 20{,}000$ and $E = 1{,}000$, therefore $I_1 = 20$:

hence the ampere turns equation gives

$$\tau_1/G = \Delta/I_1$$
$$= 100/20 = 5.$$

Now assume the ratio of primary copper space (or secondary) to the iron space as 1·20—a figure based on previous experience. That is, let

$$G/A = 1{\cdot}2.$$

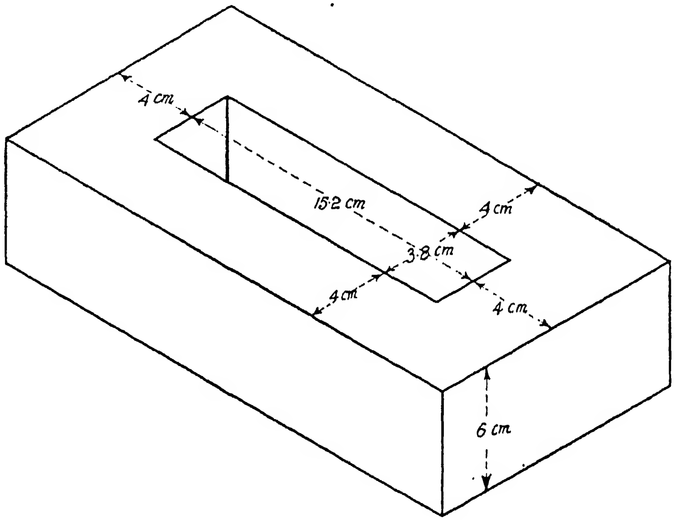


FIG. 223.

Then from the three equations we have

$$G = 29{\cdot}1$$
$$A = 24{\cdot}2$$
$$\tau_1 = 145{\cdot}4.$$

The ratio of turns is approximately equal to the ratio of voltages, therefore $\tau_2 = \tau_1 \times 10 = 1{,}454.$

We shall suppose that the core is of rectangular section and that the window is rectangular (see Fig. 223). For a preliminary design we may take the window four times as long as wide—changes in this dimension within limits having only slight effect on the regulation and the iron losses. Then, if the width be x, we have, adding primary and secondary sections

$$2G = x \times 4x$$

or $$x = \sqrt{\tfrac{1}{2}G} = 3{\cdot}81.$$

In the same way the linear proportions of the iron section are assumed—let us suppose the stampings to be y wide and piled to a depth 1·5 y. Then

$$A = 1{\cdot}5y^2$$

or

$$y = \sqrt{(A/1{\cdot}5)} = 4{\cdot}02.$$

We thus arrive at Fig. 224.

We proceed to find the gauge of wire in each winding. Since the space factor is 1/3 only $\frac{1}{3}G$ is occupied by $\tau_1 = 145$ turns. Therefore the cross section of each primary wire is $G/3\tau_1 = 0{\cdot}067$ sq. cm. The secondary turns are ten times as numerous and therefore the wire is of cross section 0·0067 sq. cm.

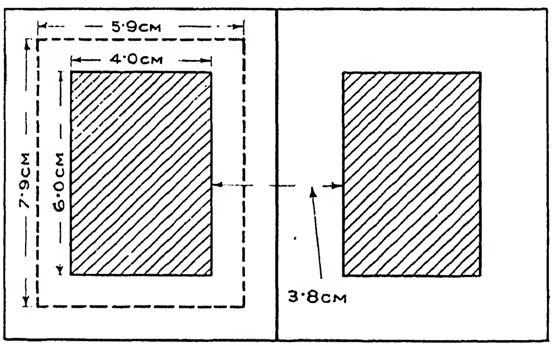


FIG. 224.

Calculation of Losses.

105. The iron losses in thin sheets of alloyed (silicon) iron may be estimated from the formula:

$$\text{Total iron losses per cycle per cubic metre in joules} = \left\{8f(\delta B'/1{,}000)^2 + 6\,(B'/1{,}000)^{1{\cdot}6}\right\},$$

δ being the thickness of the laminations in centimetres and B' the extreme flux density in the iron. Note that 1 cubic metre equals 10^6 cubic centimetres. This formula is a slight modification of a formula in G. Kapp's book on the transformer. The first term relates to the eddy loss, the second to the hysteresis loss. In ordinary transformer iron the eddy loss is about double, the hysteresis loss half as large again as that in the formula. Applying it to the dimensions of Fig. 223 we find, since there are 1,000 cycles per second, and since

$$\begin{aligned}\text{volume of iron} &= \text{volume of carcase} \times \text{space factor} \\ &= 1{,}313 \times 0{\cdot}8 = 1{,}044 \text{ cm}^3,\end{aligned}$$

therefore $$P_i = 54 + 176 = 230 \text{ watts.}$$
The copper loss is calculated from the fact that the rate of loss in watts per cubic centimetre is equal to the resistivity multiplied by the true current density provided eddy current loss in the copper can be neglected. Thus we obtain the formula

$$P_c = V_c \rho \Delta_c^2,$$

where ρ the resistivity equals $1{\cdot}56 \times 10^{-6}$ ohm/cm³, $\Delta_c = 300$ amp/cm² and V_c the volume of the copper in cubic centimetres must be calculated from Fig. 224. The mean length of a copper turn is 27·6 cm. Hence the total volume is

$$V_c = 2 \times 27{\cdot}6 \times 28{\cdot}9 \times 1/3 = 532 \text{ cu. cm.}$$

For the two coils we have the loss

$$P_c = 76 \text{ watts.}$$

Adding $$P_i = 230 \text{ watts}$$

we find that the total loss amounts to 306 watts. The calculated efficiency of this transformer is therefore nearly 98·6 per cent. The copper loss shows, by the way, that the resistance of the primary is, by Joule's equation,

$$R_1 = P_c/I_1^2 = 76 \div 20^2 = 0{\cdot}19 \text{ ohm}$$

and that the resistance of the secondary is 19·0 ohms. It should be noted that the estimate of the copper loss given above is too favourable. Even with braided stranded cable there would be appreciable eddy current loss in the copper. It is worthy of note that as a rule the best proportioned core type transformer has greater copper than iron losses while the opposite is true for the shell type transformer. The core type is better than the shell type for high voltages because the copper space factor may be made higher in the core type with the same margin of safety.

Variation of Efficiency with Load.

106. When the primary applied emf is approximately constant at all loads and the primary resistance negligible, as is nearly always assumed in power transformer design, the core flux has a fixed amplitude, and therefore the core losses are independent of the load. On the other hand, the copper losses are proportional to the square of the current (primary or secondary). Let I_2 be the load, as expressed by the secondary current, then the copper losses are proportional to I_2^2 and may be written cI_2^2 when c is a constant. Let the constant core loss be a. Also,

let us write the work done as V_2I_2, where we shall assume V_2 to be constant. Thus the total energy consumption is

$$a + V_2I_2 + cI_2^2.$$

The efficiency η is the ratio of the useful work V_2I_2 to the total energy consumption, and therefore

$$\eta = \frac{V_2I_2}{a + V_2I_2 + cI_2^2}.$$

The efficiency is therefore different at different loads, that is at different values of I_2. It is a maximum when

$$\frac{a + V_2I_2 + cI_2^2}{V_2I_2} \quad \text{or} \quad \frac{a}{V_2I_2} + 1 + \frac{cI_2}{V_2}$$

is a minimum. Now the variable part of this is of the form

$$\frac{a}{I_2} + cI_2$$

which is well-known to have its minimum value when $a = cI_2^2$. In words, the efficiency is a maximum when the core loss a is equal to the copper loss cI_2^2. To put it another way, the load that gives maximum efficiency is of magnitude

$$I_2 = (a/c)^{\frac{1}{2}}$$

and then the efficiency is

$$\eta_{max} = \frac{V_2\sqrt{a}}{2a\sqrt{c} + V_2\sqrt{a}} = \frac{V_2}{V_2 + 2(ac)^{\frac{1}{2}}}.$$

Transformer Design in Radio Circuits

Power transformer design is the very simplest case of the general problem, which deals with circuits comprising condensers as well as resistances and inductances. A full discussion of the general problem would take more space than can be afforded here and therefore we shall only touch upon the resonance adjustment, and shall do that as briefly as possible. The treatment of this problem will, however, enable the reader to form his own methods for the most general case.

107. Let the problem be to design the transformer shown at M in Fig. 225, whose function is to take current from the radio frequency machine at E and adapt it to the antenna requirements. We may suppose that the quantities given will include the frequency f, the voltage V_2 to which the antenna must be

raised, the antenna resistance R_2, the amount of inductance to be left for final tuning of the antenna, say L_b, and information about the behaviour of the machine on various non-reactive loads, including the synchronous inductance L_a and the terminal emf E at the probable load.

108. First calculate $\omega = 2\pi f$ and then the antenna current from

$$I_2 = C_2\omega V_2.$$

This leads to the power $P = R_2 I_2^2$,
whence the primary current $I_1 = P/E$.

By the ordinary rules of tuning we have $L_2 = 1/C_2\omega$.

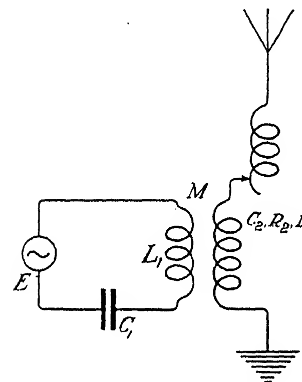


FIG. 225.

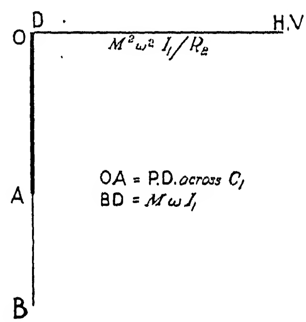


FIG. 226.

At this point we must refer to the crank diagram in Fig. 226 wherein the primary resistance is taken as negligible and M is supposed greater than L_1, which implies that the secondary turns in the transformer exceed the primary. We obtain the value of M from quantities already determined by writing the value ODH in its two possible ways, giving

$$M = R_2 I_2 / I_1 \omega.$$

In order to determine L_1 we shall suppose the leakage between the transformer turns very small and therefore

$$M^2 = (L_1 - L_a)(L_2 - L_b).$$

Having obtained L_1 we get C_1 by

$$C_1 = 1/L_1\omega^2.$$

These calculations are preliminary to the determination of the dimensions of the transformer. We shall base the method of

doing this on the principles of § 103 and we shall find that the new circumstances introduce profound modifications. As the crank diagram indicates, the secondary current is now in quadrature with the primary current instead of flowing oppositely, with the result that the " common flux " in the iron is very different from that of the non-reactive case; and, besides, the primary and secondary copper sections are in general very different. The ratio of turns in the transformer may, again, be very different from the voltage ratio V_2/E.

109. Beginning with the ratio of turns of the transformer, we see that the self-inductance of the secondary, which is proportional to τ_2^2, is $L_2 - L_0$ and therefore

$$\tau_2/\tau_1 = (L_2 - L_0)/M.$$

We now arrive at the electromagnetic equations. There will be two ampere-turns equations, the same as in § 100 except that provision must be made on this occasion for the difference between the sections of the primary and secondary windings. Let G' and G'' represent these cross sections, and let the current density multiplied by the space factor be Δ, the same in both, then since the magnetomotive effects of turns add we have

$$\tau_1 I_1 = G'\Delta \text{ and } \tau_2 I_2 = G''\Delta.$$

The final equation obtainable from the data is the common flux equation. From Fig. 226 we see that the emf BV generated by the variation of the common flux is given by

$$\begin{aligned} \mathrm{BV}^2 &= \mathrm{BD}^2 + \mathrm{DH}^2 \\ &= M^2\omega^2(I_1^2 + I_2^2). \end{aligned}$$

The common flux is of amplitude BA as B represents the maximum value allowed to the flux and A the area of section of the iron. Then as in § 103 we obtain

$$A\tau_1 = \sqrt{2}\,.\,M(I_1^2 + I_2^2)^{\frac{1}{2}}/B.$$

The square root of 2 appears in this equation because the currents are in rms measure, though B is an amplitude.

110. The three equations for τ_1/G', G'' and $A\tau_1$ are not sufficient to determine the four unknowns τ_1, G', G'', A and therefore a subsidiary condition can be satisfied. The possibilities are too various to discuss here. But their discussion would lead to, or experience would suggest, a suitable ratio for G'/A. Admitting this the problem of the preliminary design reduces to arithmetical operations like those of § 104.

Chapter V

IONIC TUBES

1. The name " ionic tube " may be taken to mean in general any tube or bulb whose chief useful functions depend upon the presence and motion in the bulb of ions—either electrons or charged matter. Ionic tubes therefore comprise, for example, X-ray bulbs, the Fleming valve, the three-electrode vacuum tubes used in wireless telegraphy, but not incandescent filament lamps. When the ions are produced mainly by the emission of electrons from heated matter in the tube the name " thermionic tube " is appropriate. The ionic tubes used in various departments of wireless telegraphy are almost all of the thermionic class at the date of writing and contain two, three or four separate conductors. They are made in all sizes ranging from 2 or 3 to more than 1,000 cu. cm in cubical content. The heated matter for supplying the electrons is usually in the form of a filament, such as is used in electric lamps, with its two ends led out of the tube to a source of current for heating. The other conductors are usually parts of circular cylinders coaxial with the filament, or plane surfaces parallel to the filament, each having its own wire through the glass of the bulb.

2. All these internal conductors are called electrodes—" ways of amber "—from the ancient association of rubbed amber and electricity. If the tube is highly exhausted and the filament cold it is found that electricity cannot be dragged across the vacuum from one electrode to another even by enormous electromotive forces; but when the filament is incandescent a current can easily be caused to traverse the vacuum from a cold electrode to the filament, though there is still difficulty in sending current in the opposite direction. The electrode by which the current enters the vacuum is called the anode; and as the current across the vacuum leaves the vacuum by entering the glowing filament this is called the cathode. The terminals of the filament are not looked upon as separate electrodes, for the principal fea-

tures of these tubes arise from the currents traversing the vacuum and not from those traversing and heating the filament. A bulb with two electrodes, namely, anode and cathode, is called a diode tube; one with three electrodes a triode, one with four electrodes a tetrode. Thus a triode tube is one with "triple ways" into the vacuum.

Thermionics.

3. The electrical phenomena described above are due to the fact that the incandescent metal of the filament emits electrons which, on account of their origin, are sometimes called thermions. They are very small pieces of negative electricity apparently all the same size. The charge on each is about $1{\cdot}56 \times 10^{-19}$ coulomb, the mass of each $8{\cdot}8 \times 10^{-28}$ gram. Their diameter has been estimated as $3{\cdot}7 \times 10^{-13}$ cm. They are very small even compared with the molecule of a gas. For example, the hydrogen molecule has been estimated to have a diameter of $2{\cdot}17 \times 10^{-8}$ cm. This is about 60,000 times larger than the diameter of an electron. Thus, if a hydrogen molecule were magnified till its diameter were the length of a large Atlantic liner, the electron would have a diameter of about 1/10th of an inch, that is to say, it would be about the size of a small raindrop. It may be remembered that it has been said that if a drop of water were magnified to the size of the earth a molecule would be about the size of a cricket ball; but such magnification would leave the electron only about 1/200th of a millimetre in diameter.

In the light of the modern theory of matter the atoms are each constituted of a positive nucleus with electrons as satellites. In a non-conductor of electricity the positive and negative portions of the atoms are bound together in an elastic manner and an electric field produces between them additional relative displacement which annuls itself when the field is removed. In a conductor, on the other hand, the smallest electric field is sufficient to move electrons throughout the substance. Apparently the positive electricity in a solid metal is permanently resident in the atoms, whereas a certain proportion of the electrons is continually skipping from one atom to another. It is in the moments of perfect liberty between two atoms that the electron can be acted upon by electric force and caused to contribute for a brief

instant to that general drift through the conductor which we call an electric current. In all cases an electrically neutral conductor has in it as much negative electricity as positive; at ordinary temperatures most of the negative electrons are at any instant under the controlling influence of the positive nuclei.

4. It has been estimated that the number of free electrons in a cubic centimetre of cold metal is about 10^{24}. The number of molecules in a cubic centimetre of copper is less than half that number. Molecules and electrons alike all partake in the irregular chattering state of motion which is the heat possessed by the body, and as the temperature is raised the chattering and the collisions increase in rapidity and violence, and the velocities of the electrons especially grow rapidly. Leaving aside for a moment the behaviour of the electrons, it is well known that if the agitations due to temperature become great enough in even a small proportion of the molecules near the surface these may attain such velocities as to be projected through the surface in spite of the inward attractions. The substance is said then to be evaporating, and, clearly, the rate of evaporation is greater the higher the temperature. Modern investigation has shown that similar considerations apply to the electrons within the conductor; the electrons near the surface that happen to attain a sufficiently high outward velocity are able to leave the conductor in spite of electrical forces tending to drag them back. Evidently the proportion attaining the necessary speed is greater the higher the temperature.

Temperature and Emission.

5. In 1903 O. W. Richardson pointed out that this view of the affair implied that the equation connecting vapour pressure with temperature should apply with appropriate modification to the analogous case of electrical evaporation. He showed that the electrical evaporation may be expressed by

$$j = aT^{\frac{1}{2}}\epsilon^{-b/T}$$

where j is the quantity of electrons escaping per second, T the absolute temperature, and a and b are constants. Richardson determined the values of the constants of several substances, and Irving Langmuir, S. Dushman, E. R. Stockle, K. K. Smith and others have contributed to our knowledge of these constants. In

Fig. 227 Langmuir's values for tungsten are written across the figure. For the same metal K. K. Smith found per square centimetre of surface the values $a = 6{\cdot}74 \times 10^8$, $b = 54{,}700$. Using vacuum tubes in which the pressure was less than one-millionth part of an atmosphere, Stockle found for molybdenum $a = 5{\cdot}0 \times 10^{11}$, $b = 8{\cdot}1 \times 10^4$.

Langmuir has recently described a coated tungsten filament

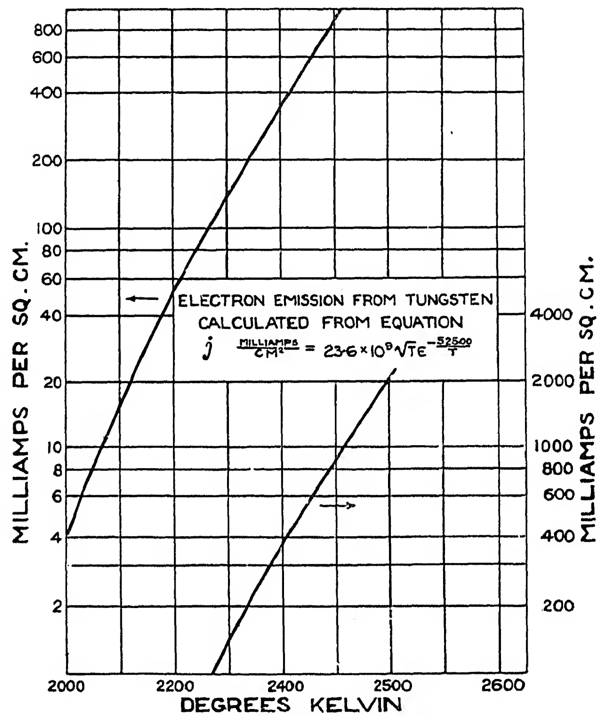


FIG. 227.—Electron Emission from Tungsten in a "Perfect Vacuum."

that emits about 3 mA of pure electron current per square centimetre at 1,300° or 1,400° K. Uncoated tungsten would have to be heated to nearly 2,000° K to give this emission. The filament is made by adding thoria to the tungsten powder before drawing and then heating the filament to a very high temperature in an extreme vacuum from which all oxygenous gases have been removed with special care. It may be that metallic thorium is

brought to the surface of the tungsten by this heat treatment and that thorium emits electrons much more freely than tungsten, or, possibly, the thoria acts in its unreduced condition.

It is evident from these data that the emission per square centimetre is different for different substances at the same temperature. According to Richardson and Langmuir, the relative emissive powers of different materials are determined by their contact differences of potential. At one time it was thought that contact difference of potential was due to chemical actions between the materials and the surrounding atmosphere; but opinion now inclines to regarding it as an intrinsic property of

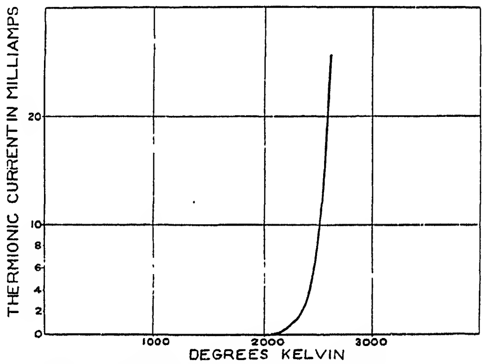


FIG. 228.—Emission from a Tungsten surface of area 0·01 sq. cm.

each material. Apart from this the emission appears to depend upon the condition of the surface of the hot conductor.

6. The connection between emission and temperature is shown strikingly in Fig. 228. On such a drawing only a very small range of temperature can be dealt with accurately owing to the rapidly growing rate of increase of the emission after the temperature of 2,000° K is passed. For this reason Fig. 227 is plotted semi-logarithmically. At 3,000° K, which is beyond the scale of that figure, the emission from tungsten is given by Langmuir as 31·7 amperes per square centimetre. Just outside such a surface the electron density is of the order 10^{20} per cubic centimetre.

From Richardson's equation in § 5 we may deduce a useful rule for calculating approximately the change in emission as the temperature varies between 2,000° K and 3,000° K, namely:

The increase in emission when the temperature rises t per cent. from T is approximately equal to bt/T per cent. of the emission at temperature T.

Motion of Electrons.

7. Since an electron is a particle of negative electricity it experiences a mechanical force when placed in an electric field and if free to move it is accelerated by the force. Experiments of many kinds, when interpreted by the reasoning of ordinary dynamics, indicate that an electron possesses a mass of about 9×10^{-28} gram. An electron gathers kinetic energy when it moves in obedience to an electric field, just as a stone acquires kinetic energy when falling under gravity. In a vacuum the motion is unresisted and therefore all the work done is converted into kinetic energy. Let v represent the difference of potential between two points on the path of the electron, q its charge, m its mass, u_1 its initial velocity and u_2 its final velocity; then the initial kinetic energy is $\frac{1}{2}mu_1^2$, the final is $\frac{1}{2}mu_2^2$, the electrical work done is qv, and, consequently,

$$qv = \tfrac{1}{2}m(u_2^2 - u_1^2).$$

As all electrons have charges and masses of the same magnitude it is possible to express the kinetic energy of an electron by the voltage that will be needed to generate that amount of energy. For example, the electrons in a metal at ordinary room temperature have an average speed of about 12×10^6 cm per second, which is about 750 miles per second; therefore we may deduce from the equation that their average speed corresponds to 0·038 volt. Again, the speed that must be attained by an electron in order that it may escape from a tungsten filament at 2,000° K is equivalent to about 4·3 volts. In order to ionise a gas by collision between molecules and flying electrons the latter must possess kinetic energy corresponding to some value between 8 and 20 volts. It is plain that the same method of stating the speeds of molecular ions, positive or negative, can be adopted in any instance where all the ions are of the same known mass and charge.

Maxwell Distribution of Velocities.

8. The value assigned in the last paragraph to the average velocity of electrons in unheated metal is arrived at by aid of

the kinetic theory of gases, the electrons being treated as if they were indeed the molecules of a special gas. By an application of the theory of probability, Maxwell deduced many years ago the law that must govern the distribution of the velocities among the molecules of an ordinary gas at various temperatures. Thus 16·1 per cent. of all the molecules have speeds between 0·9 and 1·1 times the most probable velocity, while 68·4 per cent. have speeds between 0·5 and 1·5 times the most probable velocity. Only 3·1 per cent. have velocities greater than 2·5 times the most probable velocity. On applying Maxwell's law to the electrons emitted from a heated conductor the percentage of electrons moving at any assigned speed on emission can be computed and these speeds may be conveniently recorded as voltages as already explained. For tungsten at 2,400° K 90 per cent. of the emitted electrons have speeds corresponding to above 0·022 volt, 75 per cent. to 0·059, 50 per cent. to 0·143, 25 per cent. to 0·29, 10 per cent. to 0·48, 1 per cent. to 0·95 volt. Only one electron in a thousand has the speed corresponding to above 1·42 volt, one in a million to 2·85, one in a billion to 4·27 volt. The average velocity corresponds to 0·31 volt. Numbers such as these are directly proportional to the absolute temperature. They do not depend on the material from which the electrons come.

9. One or two consequences of interest follow from this analysis of the emission. In the first place it is clear that the numbers may be read to show that if there be applied around the filament a converging electric field tending to repel the electrons back into the filament, less than 1 per cent. of the electrons will reach a distance from the filament where the electric potential is one volt above that of the metal. It is necessary to say less than 1 per cent. because not all the electrons in any speed group are emitted perpendicularly from the surface. Another consequence immediately obvious is that energy must be taken from the hot body to provide that of the electrons. By analogy with the evaporation of matter we may speak of this heat as the latent heat of the electron emission. Again, returning to the imagined experiment with the converging electric field, it is clear that it is easy to set up a field that will bring practically all the electrons to rest within a short distance and then set them into motion towards the filament. When they get near the filament they fall within the range of the forces they previously escaped from, and, in

general, enter the metal. In doing this they give back the energy they took away when they left the surface; this is analogous to the heat of condensation of a vapour. When all the electrons that leave a surface continually return to it there is as much heat returned as taken away, that is, there is thermal equilibrium in so far as the emission is concerned.

Incidence of Electrons on Matter.

10. When moving electrons impinge on the surface of any substance they come within the range of attractive forces and may thereupon end their free career. When the substance is an insulator they form a layer on the surface, thus charging it negatively. The electric field of this layer tends to prevent later electrons from reaching the surface and hence a limit may be expected to the process of accumulation of charge. When the substance is, on the other hand, a conductor, the electrons mix with the crowd of electrons normally present in such a body; and the reception of electrons will go on indefinitely if the potential is kept constant by other means. If, for instance, negative electricity is continuously drawn away by a wire the convection current in the vacuum becomes an ordinary electric current along a conductor. In either of these cases, it should be noted, the kinetic energy of the electrons absorbed is converted into heat. This bombardment of a material substance may easily be made vigorous enough to raise it to a high temperature.

11. It must be mentioned, however, that not all of the electrons are absorbed in general even by a conducting surface. Slow-moving electrons may be abundantly reflected; as large a proportion as 50 per cent. has been observed in some cases. This phenomenon recalls another one which is known as secondary emission. It is found that rapidly moving electrons on striking an unheated metal surface are capable of knocking electrons out of the molecules on this surface, and under certain suitable conditions the number so emitted from the cold surface may be many times greater than the number of primary electrons striking the surface.

Properties of a Filamentary Cathode.

12. Let us consider the case of a round filament of length l, diameter d, maintained at any constant temperature by means

of a current i and voltage e. At this temperature let the resistance per unit length of a cylinder of unit diameter be R and the thermionic emission per unit length of the unit cylinder be J. Electrical work is done in the filament at the rate ei, and this is spent in evaporating electrons, in radiation of heat and light, and in conduction to the leads. This last may be neglected at the central parts of a long filament, to which we will at present confine our attention, and, since both the other modes of spending energy are proportional to the surface area of the filament, each of them is also proportional to ei. We therefore write

$$ei = Ej,$$

where E is a constant depending on the temperature and the material of the filament, but not on the dimensions, and j is the total emission current from the length l of filament of diameter d. From the definition of J we have also

$$j = Jld.$$

In addition, by Ohm's law and the definitions above, we have

$$ei = Ri^2l/d^2 = e^2d^2/Rl.$$

In all there are three independent equations and these may be combined in various ways to give useful relations between the properties of a given filament and the constants E, J and R of the unit filament.

13. We now proceed to derive some of the most useful relations. From the first and second of the equations above we obtain the filament power equation

$$\frac{ei}{ld} = EJ,$$

while the third equation gives

$$\frac{i^2}{d^3} = \frac{EJ}{R}.$$

This last yields by aid of the first equation the relation

$$\frac{j^2}{e^2d^3} = \frac{J}{ER}.$$

The last two relations have resulted in the elimination of l; by eliminating d in a similar way we obtain the relations

$$\frac{j^3}{i^2l^3} = \frac{J^2R}{E}$$

and
$$\frac{e^2 j}{l^3} = EJ^2R.$$

Further, two other equations with three variables may be expected; these are
$$\frac{e^2 d}{l^2} = EJR$$

and
$$\frac{e^3 i}{l^3} = E^2J^2R.$$

14. The preceding seven relations have been so arranged as to show at once that the functions of voltages, currents and lengths on the left hand are constant at constant temperature. In other words, if measurements be made on any filament and these functions on the left hand computed, the constants of the unit cylinder are obtained variously combined as indicated. The data have been obtained, discussed and recorded by Langmuir, Dushman, and more recently by G. Stead. From their results values of J and R may be derived. These depend on the nature of the sample of metal under investigation. The quantity E, indeed, depends also upon the character of the surroundings; for example, it is smaller when the filament is inside a reflecting shield than when it radiates heat and light freely. For commercially pure tungsten the approximate values are

	2,100° K.	2,300° K.	2,500° K.	2,700° K.
E	2,180·0	298·0	57·8	24·0
J	0·060	0·68	5·21	17·0
R	$8{\cdot}3 \times 10^{-5}$	$9{\cdot}4 \times 10^{-5}$	$10{\cdot}5 \times 10^{-5}$	$11{\cdot}7 \times 10^{-5}$

In the above table J is the emission of the unit cylinder in amperes and R is the resistance of the unit cylinder in ohms. The quantity E is of special interest because of its appearance in the first equation of § 12; for clearly $1/E$ gives the emission in amperes per watt of heating current for all lengths and diameters of filament, and the table shows it increases with the temperature.

Calculation of Filament Dimensions.

15. The equations developed above can now be used for practical calculations concerning the dimensions of filaments, always supposing the effects of the cooling of the ends of the filament by conduction to the supports can be neglected. The equation for i^2/d^3 is a notable one—it gives the filament current in terms of the diameter for all values of the length, the emission and the terminal voltage. But of course the equations required in solving any problem must be selected from the above for the occasion. As an example, let it be required to find the emission, the filament current and the length of the filament which could be used directly on a 10-volt battery at a temperature of 2,300° K, the diameter of the filament being given as 0·02 cm. We may use the EJ/R equation for obtaining the filament current, the J/ER equation for the emission and the EJR equation for the length. At 2,300° the quantities just enumerated have the values $2{\cdot}13 \times 10^6$, $2{\cdot}46$ and $1{\cdot}88 \times 10^{-2}$ respectively. We find

$$i^2 = 2{\cdot}13 \times 10^6 \times d^3,$$

or filament current $i = 4{\cdot}13$ A.

Again $j^2 = 24{\cdot}6 \times e^2 d^3,$

or emission $j = 0{\cdot}14$ A.

Also $l^2 = e^2 d \div (1{\cdot}88 \times 10^{-2}),$

or length $l = 10{\cdot}3$ cm.

16. The end correction for cooling in a case like this is probably less than 10 per cent. even in the value of the emission; it is greater for shorter filaments. Langmuir states that under ordinary conditions the cooling at the two ends of a tungsten filament lowers the uncorrected electron emission by an amount corresponding roughly with the emission from a length of filament in which the voltage drop is 1·4 volts. Useful empirical methods of making the correction have been developed by G. Stead in the *Journal of the Institution of Electrical Engineers*, January, 1920.

DIODES

17. Let us consider the electrical phenomena that arise in a tube containing a hot electrode and a cold one in consequence of the emission of electrons from the hot electrode. The simplest case for our purpose will be that in which both electrodes are

plane and placed so near together that their distance apart is small compared with their linear dimensions, but this form is never used in practice on account of the difficulty of maintaining a large plane electrode at a uniform high temperature. We shall therefore take the less simple case of a cold electrode consisting of a hollow circular cylinder with a hot electrode consisting of a filament lying along the axis of the cylinder. These electrodes will be supported in a glass tube in some such manner as is indicated in Fig. 229, and the glass vessel will be supposed perfectly evacuated. We shall assume that both electrodes are of tungsten. The filament must be heated by passing a current through it, but we shall ignore the other effects of this current for the present and particularly those due to the fall of potential along it when the heating current flows.

18. In order to fix ideas the cylinder may be supposed to be 2 in. long and 1 in. in diameter, and the filament may be taken to be 5 mils in diameter. Let us start with the filament heated to a temperature of 1,320° K and let the anode be connected directly to one end of the filament through a sensitive galvanometer. It will be found that a current of 10^{-9} of an ampere will pass through the galvanometer from the filament to the cylinder. This is due to the fact that the electrons emitted from the filament with various velocities, as described in § 8, succeed in crossing from the filament to the cylinder in large numbers and cause a flow of the electrons in the connecting wires of the windings of the galvanometer from plate to filament which we regard, in obedience to the convention fixed long ago, as a positive current in the opposite direction. This small current produces practically no difference of potential between the cylinder and the filament, and therefore practically all the electrons emitted from the filament cross the vacuum and enter the cylinder. If now a battery be inserted in series with the galvanometer so as to make the cylinder positive with respect to the filament, the deflection of the galvanometer will not be increased, for the reason that all the

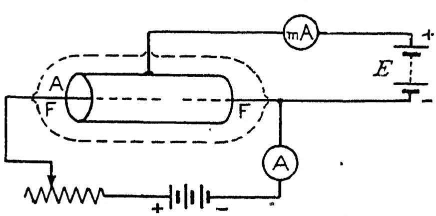


Fig. 229.—A diode and connections.

electrons emitted from the filament were already being taken to the cylinder. The applied voltage may be changed from 1 volt to 100 volts without altering the value of the current flowing. This is therefore called the saturation current.

19. Now suppose the cylinder be given a small negative potential with respect to the filament, say by inserting a voltaic cell, it will be found that the current diminishes. This is because all those electrons emitted with a velocity less than that corresponding to the applied negative voltage are repelled into the filament; in fact the proportional reduction of the current is determined by the Maxwell law of distribution of velocities, and

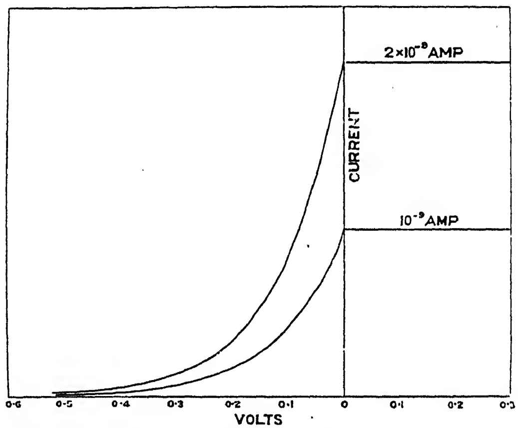


FIG. 230.—Current Curves at two low temperatures of the cathode.

therefore the current corresponding to each value can be determined from Maxwell's formula. For example, in the present case 10 per cent. of the electrons can flow to the anode against a reverse emf of 0·26 volt and 1 per cent. against 0·52 volt, the current being 10^{-10} ampere and 10^{-11} ampere in these cases. The application of the Maxwell law is different with different forms of electrodes because the component rather than the total velocities have to be considered, but it is always of exponential form, and therefore the currents observed in our present experiment are expected to decrease exponentially as the applied negative voltage is increased. This is indicated in Fig. 230. The chief arithmetical property of such curves is that the ordinates change

in geometrical progression when the abscissæ vary in arithmetical progression. In the curves here shown the current increases tenfold when the emf is increased by 0·26 volt at any point. From these considerations it is evident that by increasing the negative potential of the cylinder the electron current flowing can be reduced until at last it cannot be detected, although theoretically it is never completely stopped.

20. Now let us increase the temperature of the filament and return to the case when there is no emf applied externally. The current through the galvanometer may be expected to increase according to the law enunciated in § 5. Suppose, for instance, that the temperature be raised sufficiently to double the current, then we may calculate the new temperature by aid of the equation in § 5 or by aid of the approximate equation

$$T_2 - T_1 = \left(\frac{T_1}{235}\right)^2,$$

which is valid for small changes of temperature. We obtain the result 1,343° K. If a positive potential be given to the cylinder the saturated current is obtained as in the previous experiment. This is indicated by the upper level line in Fig. 230.

In order to find the shape of the current curve when negative potentials are given to the cylinder it is only necessary to notice that the distribution of velocities is not affected by the temperature change and that therefore every ordinate of the former current curve must be doubled to represent graphically the results of the present experiment.

Space Charge.

21. Let the filament now be heated to a temperature of about 2,400° K. At this temperature the emission of electrons amounts to 50 milliamperes. This is fifty million times greater than in the experiment dealt with in the last paragraph. The galvanometer must therefore be replaced with a milliammeter. If there be no source of emf in circuit it will be found that the milliammeter shows no deflection. No doubt if a delicate instrument were placed in circuit a larger current would be observed than was obtained at the lower temperatures, but the striking fact is that the number of electrons which manage to cross the space when the filament is at a high temperature

is a very small fraction of the total number emitted. The rules developed in the preceding paragraph do not apply now.

If now a battery be inserted so as to make the cylinder positive relative to the filament, electrons are attracted across the vacuum and the current is found to increase faster than in proportion to the applied voltage. When this reaches the value of 90 volts the current has become 50 milliamperes ; if the voltage be raised still further the current remains constant, that is to say, the saturation current is being obtained because all the electrons emitted from the filament are being taken to the cylinder. Evidently between zero and 90 volts considerations arise other than that of the initial velocities of the electrons. In fact the relatively small

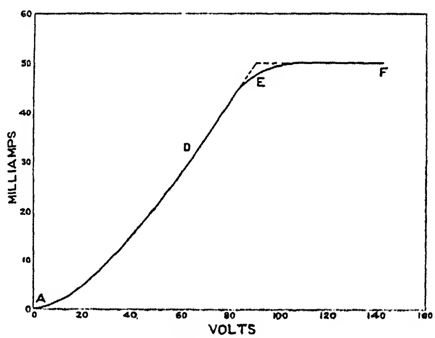


FIG. 231.

value of the current at zero voltage shows that the effects of the initial velocities are negligible, and that the electrons may for the present purpose be regarded as being emitted without velocity.

22. The new consideration arises out of the mutual repulsions between the electrons travelling across the space. Clearly, whatever the distribution of the electrons between the electrodes, those in motion across the vacuum tend to repel those near the filament, and this is stronger the greater the number of electrons in motion, that is to say, the greater the current. It can be shown that when the filament is at 1,320° K this electric repelling force is less than 0·001 volt, and, because of its insignificance, can be neglected in comparison with the mean velocity of emission ; but we have seen that with the filament at 2,400° K a potential difference of 90 volts between cylinder and filament is required in order to cancel the repulsion completely and to ensure that none

of the emitted electrons re-enter the filament. At first sight it would seem that all the electrons in the space between the electrodes would rush towards the cylinder however small the potential difference applied, but reflection shows that the electric field of a slightly positive cylinder neutralises the repelling field of the space charge merely to a definite distance inward from the cylinder. Electrons within this distance are accelerated rapidly towards the cylinder, but those nearer the filament experience an inward push and may return to it.

The curve showing the relation between the current and the

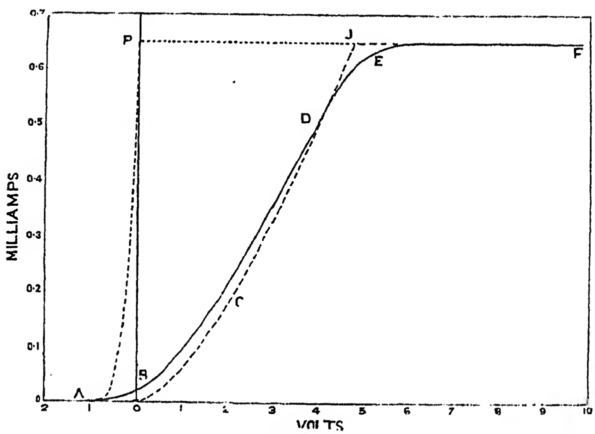


FIG. 232

applied emf, often called the characteristic curve, is shown in Fig. 231. On the scale of this curve the current flowing at zero voltage, and due to the initial velocities of the electrons, is invisible. Fig. 232 gives the curve for the case when the filament is at the temperature of 1,980° K, where the saturation current is 0·65 milliampere. On the scale of this figure the effect of the initial velocities is perceptible. The dotted curve shows the current that would have been expected if the mutual repulsion between the electrons had not been taken into account.

Theory of Space Charge

The more accurate study of the properties of the space charge demands the use of mathematics. In order to keep this as simple

as possible let us solve the problem of the space charge between two parallel planes and after that consider briefly the problem of cylinders and filament.

Cathode and Anode both Planes.

23. A sectional view is given in Fig. 233, the distance between the planes being taken as x_a and the distance of a typical plane layer from the cathode being x. We shall suppose the hot cathode to be maintained at zero potential and the plate to be maintained at a potential e_a above the cathode. The surface density of positive electricity on the anode will be indicated by σ_a. The electric force in the typical layer will be F directed from anode to cathode, and therefore the mechanical force on an electron in that layer will be towards the anode and of magnitude Fq, where q is the magnitude of the negative charge of the electron. When a steady state has been attained a current i_1 will be crossing each square centimetre of every plane between the anode and cathode and parallel to them.

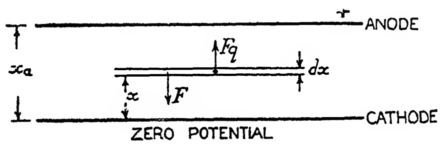


FIG. 233.—Sectional view of space between two plane parallel electrodes.

Let the number of electrons per cubic centimetre in the typical layer be n and the velocity of the electrons towards the anode be u. Then since the current from the square centimetre of anode is equal to the amount of negative electricity passing through the square centimetre of the typical layer towards the anode we shall have

$$i_1 = nqu.$$

We now have to introduce the mathematical statement of the mechanical law governing the motion. This is most conveniently done by observing that the kinetic energy gained by an electron since it left the cathode and travelled a distance x towards the anode is equal to the work done upon it by the electric forces, or, in other words, is equal to its charge multiplied by the potential difference through which it has risen. Let v be the potential at the distance x from the plate and m be the mass of the electron. Then the above statement becomes

$$\tfrac{1}{2}mu^2 = qv.$$

In forming this equation we have tacitly assumed that the velocity with which the electrons leave the hot cathode is negligible.

24. Next we must introduce a statement of the fact that the change of the electric force from point to point is due to the electricity distributed in the space. The equation expressing this in the most suitable manner is obtained by the assistance of the special diagram of Fig. 234, which represents the charge in an infinitesimal layer of thickness dx taken aside for the purpose of deducing the new equation. In the figure it is supposed that the electricity is receiving Faraday lines in different numbers per square centimetre above and below the layer. These lines, which are all directed down the page, are fewer below the layer because some have ended on negative electricity there. In passing from a point P on one side to a point P′ on the other side of the layer the change in the number of lines is equal to the number of units of negative electricity per square centimetre of the layer, which is $nqdx$. Now the electric force at any point of a field is equal to the Faraday line density multiplied by $4\pi/\kappa$. Thus the change in the downward directed field in passing from P to P′ is $(4\pi nq/\kappa)dx$. This change of field may be written dF. We obtain therefore the electric field equation

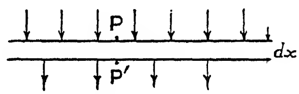


FIG. 234.—Change of electric force in crossing a layer of electricity.

$$\frac{dF}{dx} = \frac{4\pi}{\kappa} nq.$$

This and the equations already obtained, namely,

$$i_1 = nqu,$$
$$\tfrac{1}{2}mu^2 = qv,$$

together with

$$F = dv/dx$$

form a system to be solved simultaneously. It will be noticed that because the positive direction of F has been taken opposite to the direction of increase of x, the minus signs which usually appear in front of the differential coefficients are absent from the equations. This is very convenient for the present problem, where we are dealing throughout with negative electricity.

25. Before attempting to solve the equations we must notice that since we have taken the velocity of the electrons to be negli-

gible when emitted from the cathode, the electric field at the surface of the cathode must consequently be zero unless the saturation current is being taken. For if we suppose the current to be well below the saturation value, then more electrons are being emitted than are carried to the anode and the balance is continually being returned to the cathode. This action takes place because the smallest accumulation of unwanted electrons leads immediately to the creation near the cathode of an electric field which pushes newly emitted electrons back again; consequently the accumulation of free electrons is automatically prevented and the electric field has no opportunity of becoming different from zero. If, however, the applied voltage between the plates is greater than that needed to produce the saturation current there will exist at the cathode a positive field that could carry away more electrons than are actually being emitted. The condition of zero field at the cathode, and also that of zero voltage, must be satisfied by any solution we may obtain.

26. We may assume as a trial solution

$$v = ax^b,$$

where a and b are to be determined. This gives

$$F = abx^{b-1}$$

Both v and F vanish at the cathode as is required. Again, by further differentiation,

$$\frac{4\pi}{\kappa} nq = ab(b-1)x^{b-2}$$

therefore
$$u = \frac{4\pi i_1}{\kappa} \frac{x^{2-b}}{ab(b-1)}$$

and
$$v = \frac{m}{2q}\left(\frac{4\pi i_1}{\kappa}\right)^2 \left(\frac{x^{2-b}}{ab(b-1)}\right)^2.$$

But the value of v just obtained must be equal to that from which we started if the latter is really a solution. By equating the indices of x we have

$$b = 4 - 2b$$

or
$$b = 4/3.$$

By equating the coefficients we have

$$a = \frac{m}{2q}\left(\frac{4\pi i_1}{\kappa ab(b-1)}\right)^2$$

or $$a = (i_1/A)^{\frac{2}{3}},$$
if, for brevity, we put

$$A = \frac{\kappa}{9\pi}\sqrt{\frac{2q}{m}}.$$

The solution may therefore be written

$$v = (i_1/A)^{\frac{2}{3}}\, x^{\frac{4}{3}}$$

or $$A^2 v^3 = i_1^{\,2} x^4$$

or $$i_1 = A v^{\frac{3}{2}}/x^2.$$

27. This solution has been given by C. D. Child in the *Physical Review*, p. 498, 1911, and by I. Langmuir in the same Journal, p. 450, 1913. We may regard the v, i_1, x equation as showing the voltage at which an anode placed at the distance x from the cathode must be maintained in order to hold the unsaturated current i_1 constant; or we may say that when we observe a current of density i_1 being carried by electrons between two parallel planes distant x apart we know the voltage between the planes must be as given by the above equation. From the v, i_1, x equation we may obviously obtain the distribution of F, n, and u by returning to the equations for these quantities and substituting for a and b the values discovered in the course of the solution by trial. We find

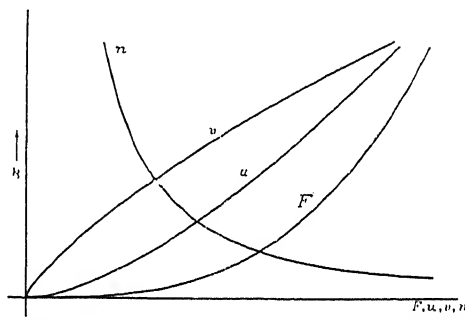


FIG. 235.—Potential and field, electron density and velocity in space between two parallel planes.

$$F = \frac{4}{3}\left(\frac{i_1}{A}\right)^{\frac{2}{3}} x^{\frac{1}{3}}$$

$$nq = \frac{\kappa}{9\pi}\left(\frac{i_1}{A}\right)^{\frac{2}{3}} x^{-\frac{2}{3}}$$

$$u = \frac{9\pi}{\kappa} A^{\frac{2}{3}} i_1^{\,\frac{1}{3}} x^{\frac{2}{3}}.$$

The distribution of these quantities between the electrodes is given by the curves of Fig. 235, which are not drawn to scale. The

total charge in motion at any time between unit area of the plates is obtained by integrating nq with respect to x; it is

$$\frac{\kappa}{3\pi}\left(\frac{i_1}{A}\right)^{\frac{2}{3}} a^{\frac{1}{3}},$$

and is proportional to the value of the field at the anode This is directly evident from the consideration that the Faraday lines leaving unit area of the anode are $\kappa F/4\pi$ in number, and these are all occupied in reaching out, so to speak, to the negative electricity in the space between the plates.

The charge of the electron is known to be $q = 1{\cdot}56 \times 10^{-20}$ emu $= 1{\cdot}56 \times 10^{-19}$ in practical units; the mass of the electron is $m = 8{\cdot}8 \times 10^{-28}$ g $= 8{\cdot}8 \times 10^{-35}$ practical units; therefore $q/m = 1{\cdot}77 \times 10^{15}$ practical units. Also $\kappa = 10^{-11}/9$ practical units. Hence

$$A = 2{\cdot}33 \times 10^{-6} \text{ in practical units.}$$

As an example, let i_1 be one milliampere and x one centimetre; then if the current is not saturated the following values will be found to hold at the anode: $v_a = 57$ volts, $F = 76$ volts per centimetre, $nq = 2{\cdot}23 \times 10^{-12}$ coulomb per cubic centimetre, $u = 4{\cdot}47 \times 10^8$ centimetres per second.

Effect of Initial Velocities.

28. In the preceding paragraphs the electrons are imagined to emerge from the cathode surface with negligible velocity. Let us now remove this limitation and let us assume that they all leave the cathode with the same normal velocity u_1. If the current is not saturated a certain number of electrons must be returned to the cathode; these electrons are stopped and turned back by virtue of a negative electric field, and this field must be due to an accumulation of electrons at a certain distance from the cathode. At this place the field must be zero (for accumulation to be possible); and as F changes sign in passing through the zero value, some of the electrons reaching and transiently resting at the place of accumulation will be dragged towards the anode, and some will be sent back to the cathode. Since F is the gradient of v it is evident that v is a minimum at this place of zero velocity; we shall call the co-ordinate of this place x_0. Moreover, as the differential equations must still be obeyed, the changes in the old solution must be confined to the introduction of arbitrary con-

stants at permissible places. Inspection of the differentiations shows that we may write

$$v + v_0 = (i_1/A)^{\frac{2}{3}} (x - x_0)^{\frac{4}{3}}$$

whence

$$F = \frac{4}{3}\left(\frac{i_1}{A}\right)^{\frac{2}{3}} (x - x_0)^{\frac{1}{3}}.$$

In these equations we have, as we should,

$$v = -v_0 \text{ at } x = x_0$$

and

$$F = 0 \text{ at } x = x_0,$$

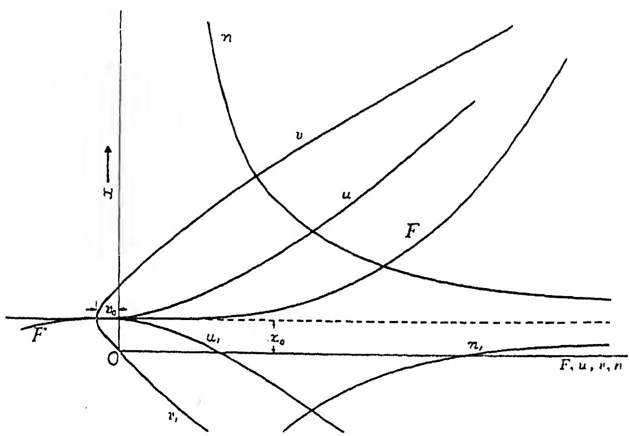


FIG. 236.—Space charge conditions when electrons emitted at velocity u_1.

while F is negative when $x < x_0$ and positive when $x > x_0$. Continuing the differentiations as before we obtain

$$nq = \frac{\kappa}{9\pi}\left(\frac{i_1}{A}\right)^{\frac{2}{3}} (x - x_0)^{-\frac{2}{3}}$$

and

$$u = \frac{9\pi}{\kappa} A^{\frac{2}{3}}\, i_1^{\frac{1}{3}} (x - x_0)^{\frac{2}{3}}.$$

When i_1 is held constant we may represent the state of affairs between the electrodes by the curves of Fig. 236.

29. The curves may be interpreted as follows: An electron leaving the cathode and moving up the page finds itself in a retarding field by which its initial velocity u_1 is reduced to zero by the time it arrives at x_0. Here the electric field changes

sign and the potential reaches its minimum value v_0. A consequence of the slowing down of the electrons is that their density n increases to a very large value at x_0, the density multiplied by the velocity being, of course, constant because the current constituted by their motion is constant. Some of these accumulated electrons may be assumed to return sooner or later to the cathode, while some pass onwards under the influence of the positive field and rising potential; and beyond x_0 the electrons gather speed and become less closely packed, as is seen from the curves of u and n.

We may estimate the magnitude of x_0 by putting v and x zero in the voltage equation above. We thus obtain

$$x_0{}^2 i_1 = A v_0{}^{\frac{3}{2}}.$$

In practical units $A = 2{\cdot}33 \times 10^{-6}$, and it is known that the mean velocity of the electrons emitted from a cathode at a temperature of 2,000° K corresponds to about 0·25 volt. Let us assume that i_1 is, in a particular experiment, of magnitude 2 milliamperes. Then

$$\begin{aligned} x_0 &= \left(\frac{2{\cdot}33 \times 10^{-6} \times (0{\cdot}25)^{\frac{3}{2}}}{2 \times 10^{-3}}\right)^{\frac{1}{2}} \\ &= 1{\cdot}2 \times 10^{-2} \text{ cm.} \end{aligned}$$

The distance x_0, which is of order a tenth of a millimetre, may therefore be regarded as small compared with the dimensions of any ordinary thermionic tube. Closer calculation indicates that this is likely to be rather too large than too small; but using the numbers just obtained we have

$$\begin{aligned} v + 0{\cdot}25 &= 74\,(x - 0{\cdot}012)^{\frac{4}{3}} \\ F &= 99\,(x - 0{\cdot}012)^{\frac{1}{3}} \end{aligned}$$

when $i_1 = 2$ milliamperes. It should be noticed that x_0 is smaller the greater the current.

30. The formation of a place of accumulation of electrons, and of a place of zero field, at a short distance x_0 from the cathode, reduces the case of electrons with finite initial velocity to that of §§ 25, 26, 27, in which the initial velocity was taken to be zero. In every case, then, we may sum up the matter by saying that when the voltage and the current carried by the space charge are small enough there is a place of zero field; an increase of voltage calls up an increase of current and a corresponding increase of the number of moving electrons, and the electric force due to this

new distribution of the charge is just right to restore a layer of zero field. If there are not sufficient electrons to restore zero field the potential gradient becomes positive everywhere and the current is saturated.

31. It has been assumed in the above discussion that the normal velocity of emission of the electrons has the same value u_1 for every electron. This is not true, for the evaporated electrons have in fact a large range of velocities. The result is that instead of a layer of accumulated electrons of infinitesimal thickness and infinite density at a definite distance x_0 we expect in fact a vaguely defined region of finite thickness and finite density in which sufficient electrons accumulate to produce the negative electric field needed for repelling the unwanted electrons back into the cathode.

Cathode and Anode Concentric Cylinders.

32. When the electrodes are of the form discussed in §§ 17 to 22 and illustrated in Fig. 229, analysis similar to that employed in the last few paragraphs may be applied to find the distribution of the space charge, but a complete discussion demands more mathematics than is desirable in this book. Readers may refer to I. Langmuir's paper in the *Physical Review* of November and December, 1913, where it is shown that when the radius of the filament is less than a tenth of the radius of the cylindrical anode, the current i' per centimetre of the length of the concentric electrodes tends towards a lowest value, given by

$$i' = \frac{2\sqrt{2}}{9}\sqrt{\frac{q}{m}} \cdot \kappa \frac{v^{\frac{3}{2}}}{x},$$

as the ratio of the filament radius to the anode radius becomes smaller. Here x is the distance from the axis of the filament to any point in the space charge and v the potential at the point above that of the filament.

The current density at any point distant x from the axis of the filament is given by

$$i_1 = \frac{i'}{2\pi x} = \frac{\sqrt{2}}{9\pi}\sqrt{\frac{q}{m}} \cdot \kappa \frac{v^{\frac{3}{2}}}{x^2}.$$

In practical units the formulæ become

$$i' = 14{\cdot}65 \times 10^{-6}\,\frac{v^{\frac{3}{2}}}{x}$$

$$i_1 = 2{\cdot}33 \times 10^{-6}\,\frac{v^{\frac{3}{2}}}{x^2}.$$

If we wish we may take x to be the radius of the cylindrical anode; and if this is not very great compared with the radius of the cathode a correcting factor must be applied to both the last formulæ. The multiplier depends upon the ratio of the cathode radius $\frac{1}{2}d$ to the anode radius x in the following manner :—

$d/2x$ . .	0·067	0·10	0·15	0·2	0·3	0·4	0·5	0·6	0·7	0·8	1·0
Multiplier .	1·02	1·06	1·18	1·32	1·75	2·41	3·61	6·00	10·9	23·3	∞

This table may be applied in order to make rough allowances for the case when, for instance, the cathode is in the form of a curly filament, or is a bundle of parallel filaments.

Extensions of Space Charge Theory.

33. Langmuir has given theoretical and experimental evidence to show that the electron current passing across a perfect vacuum between two electrodes of any shape is proportional to the three halves power of the applied voltage provided that every portion of the cathode surface is emitting a surplus of electrons and that every other conductor in the space is at the potential of either the cathode or the anode. In this statement it is assumed also that the voltages being applied are so great that the initial velocities of the electrons are negligible. The law holds over wide ranges of voltages even when the electrodes are both in the form of wires twisted into irregular shapes. Much of the reasoning concerning plane and cylindrical electrodes can therefore be extended to electrodes of any shape and position, provided that the cathode is at a sufficiently high temperature. In particular it is clear from the preceding paragraphs that if the cathode is hot enough to supply easily the current taken to the anode by any given voltage it does not matter how much hotter it may be made. In other words, the ability of the vacuous region for conveying current under a stated voltage is independent of the temperature of the cathode if every portion is hot enough to emit a surplus of electrons, and depends only upon the configuration of the electrode surfaces. The conception of the " current carrying capacity " of a space has therefore been introduced by Langmuir and Dushman, and it has been defined as the current passing through the

given tube when an emf of 10 volts is applied between anode and cathode. For many of our purposes it is more convenient to use the word " perveance " for the capability of the tube for passing current and to define the perveance at a given anode voltage as the current that will ideally be sent through the tube by an emf of one volt on the assumption that the three halves power law is followed at all voltages. The perveance is then measured in mhos, and its reciprocal, the imperveance, in ohms.

34. Let G represent the perveance of the space between two electrodes, then the current flowing when a voltage e is applied is $Ge^{\frac{3}{2}}$ amperes. The quantity G is a function of the geometry of the electrodes and the properties of the electron.

For a pair of parallel equal plates area S placed a small distance x apart

$$G = 2{\cdot}33 \times 10^{-6} S/x^2.$$

For a length l of a filament and co-axial cylinder

$$G = 14{\cdot}65 \times 10^{-6} l/x$$

where x is the radius of the cylinder. These results are merely approximate because end and edge effects are left out of account. From the latter equation we see that the perveance of all cylindrical diodes having the same linear proportions is approximately the same. For example, the perveance of the tube of § 18 with a cylinder 2 in. long and 1 in. in diameter is $58{\cdot}6 \times 10^{-6}$ mho, and that of a diode with a cylinder 2 ft. by 1 ft. is the same. An emf of 100 V can produce a current of 58·6 mA in either tube if the filament is hot enough and no more than this however hot it may be made. The " resistance " of this tube may be said to be $100 \div 0{\cdot}0586$ ohms.

It may be asked : Is work done and heat developed according to the law of Joule in forcing current through this resistance ? The answer is that, in the case cited, work is spent on speeding the electrons from the hot electrode to the cold and their energy of motion is in the broad sense a form of heat. When they strike the anode they give up their kinetic energy to its molecules and the electrons in it, and the heat then appears in the familiar form. We may therefore correctly apply Joule's law to the calculation of the heat given to the anode by bombardment by electrons. In the above instance the heat developed per second is the equivalent of $100 \times 0{\cdot}0586$ or 5·86 watts. In general, the rate of heating

is equal to the product of the voltage and the electron current. This last statement is true it may be noted even when the current is saturated and when "resistance of the tube" has no useful meaning.

35. We are now in a position to appreciate that there are two possible modes of limitation of the electron current passing from cathode to anode across a perfect vacuum. The one may be called the emission limitation, the other the space-charge limitation. The former plainly appears when the cathode is kept at a constant temperature and the voltage gradually raised; at a voltage called the saturation voltage all the electrons produced

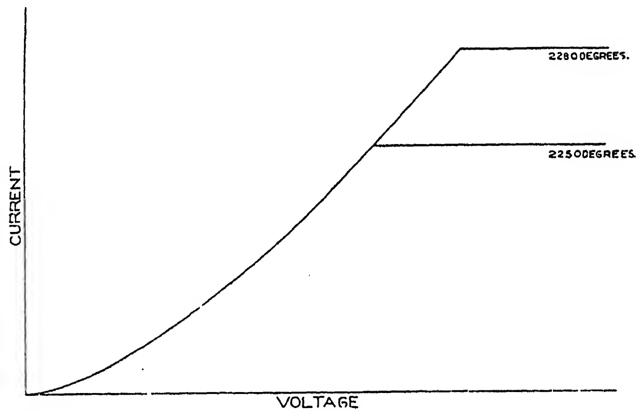


FIG. 237.

at the given temperature are taken to the anode, and higher voltages can do no more. The space-charge limitation, on the other hand, is seen when the temperature of the filament is varied and the voltage kept constant at a medium value. Starting with the filament at a red heat relatively few electrons will be emitted and these will all be carried away by the field created by the battery. The temperature may even be raised to white heat and all the electrons still be taken across to the cylinder if the voltage is large enough. In fact, the evaporation equation connecting j and T is followed accurately. Proceeding to a higher temperature an instant is reached rather suddenly when the emission becomes too copious for all the electrons to be swept up by the existing electric field and therefore the space charge begins to accumulate. After this the filament temperature may be in-

creased to destruction but there is no increase of the electron current to the cylinder. A graph of these experiments is given in Fig. 237. The curved portion follows the law expressed by the evaporation equation. Some experimental curves of Langmuir's for many values of the fixed voltage for a particular tube are given in Fig. 238. A line parallel to the current axis shows how

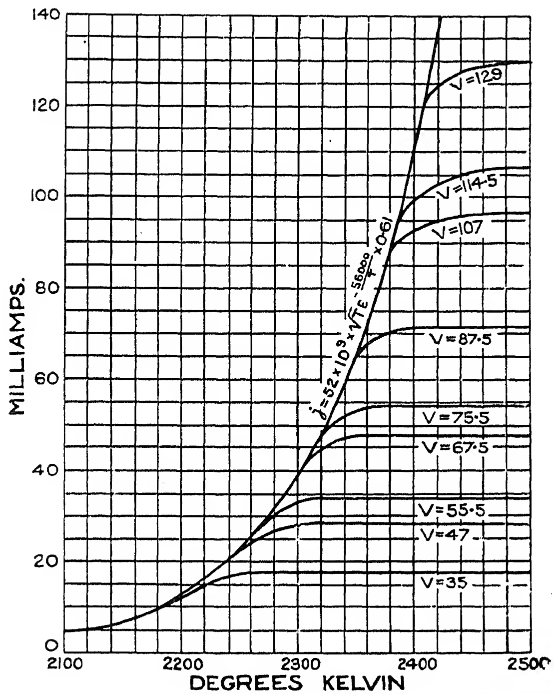


FIG. 238.—Emission and space charge limitations of the thermionic currents.

the current and voltage vary together if the temperature be kept fixed, that is, this line exhibits the three halves power law.

Deficiencies in the above Theory.

36. In the more highly exhausted forms of ionic tube met with in practice the current and voltage obey the equations of the preceding paragraphs fairly closely within certain ranges of values of filament temperature and anode voltage. But even in the most perfectly evacuated tubes departures from the equations

are to be expected and are, indeed, sometimes of utility. We proceed to review the causes of some of these occurrences.

Initial Velocities.

In the first place, we have seen in § 18 that if the filament emission is very feeble the current is wholly due to the velocities of emission of the electrons. We may call this the Maxwell current characteristic. When the emission is increased to a medium value as indicated in Fig. 232 we expect to find that the characteristic has an exponential equation under negative voltages, and that up to about 5 volts positive, when the current is partly due to initial velocities and partly governed by the space charge, the characteristic will go through a transition region where the ordinates are neither exponential nor three halves power functions of the abscissæ. At rather higher voltages, according to the discussion in the preceding pages, we may expect that the three halves power is closely followed. We may therefore call this the Langmuir portion of the characteristic.

The relative magnitude of the Maxwell portion to the Langmuir portion depends, however, to some extent on the shape and position of the electrodes; it is less marked when the electrodes are parallel planes than when they are cylindrical. This is because the convergence of the Faraday lines towards the filament in the latter case concentrates the potential drop near to the cathode and tends to make insignificant the voltages corresponding to the initial velocities. Again, the presence of another conductor such as the grid of a triode, connected to the cathode, tends to exaggerate the Maxwell part of the characteristic because it shields the cathode from the anode and reduces the potential drop near the cathode; simultaneously it cuts down the perveance of the tube.

Contact Potential.

37. When anode and cathode are of different metals, as is often the case, the "contact difference of potential" exercises an influence; it may be of magnitude as great as two volts, and therefore when voltages are low it has a perceptible effect on both the Maxwell and the Langmuir portions. Clearly therefore it may affect the Langmuir curve more than the initial velocities do. A thoriated filament or a lime-coated filament (Wehnelt cathode), when connected to a metal surface near it, tends to become posi-

tive relative to the metal because of its greater emissive power, that is, the metal tends to become relatively negative. A large area of metal near such a filament therefore tends to repel electrons into the filament and influences the Langmuir curve. The effect could be corrected by subtracting a function of the contact difference of potential from the applied voltage before raising the voltage to the three halves power.

Practical tubes do not always possess the ideally perfect filament so far discussed. Filaments to which thoria has been added in the process of manufacture may have on their surface specks of active, surrounded by less active, matter. In consequence, when the anode voltage is raised gradually, the less emissive parts of the filament become saturated before the more emissive parts and this causes the Langmuir portion of the characteristic to bend very gradually into the level saturation part. Again, in practical tubes the surface of the anode is sometimes not perfectly clean and this introduces an unexpected contact potential difference amounting to a volt or more which produces the same effects as those described in the last paragraph.

Structural Considerations

38. Mechanical defects have their influence on the characteristic. If some parts of the filament are nearer to the anode than others those parts are saturated first when the voltage is steadily raised and this causes the Langmuir portion to turn towards the saturated portion at an early stage. Electrodes of complicated shape may clearly produce departures from the three halves power curve for the same reason. The sagging of a filament when very hot may cause a tube to show greater departures at high than at low temperatures.

39. When the electrodes are so designed and spaced that the electric field between them carries electrons, or allows electrons to be carried, to the glass walls of the tube, the charging up of the glass may cause very puzzling effects. As a rule such charging brings about a reduction of the perveance and may in extreme cases reduce it to nearly zero. Langmuir describes an experiment in which the tube was brought into action by raising the filament temperature after applying the anode voltage, and then the anode current was momentarily interrupted ; with the result that the anode current was found to be much greater than before

and the characteristic completely altered. In the first part of the experiment the glass became charged negatively and this charge augmented the usual space charge repulsions under the cathode. In the second part of the experiment, the sudden application of the high voltage caused high velocity electrons to strike the glass violently and produce secondary emission of low velocity electrons. A stationary adjustment was quickly reached in which the potential of the glass rose to nearly that of the anode and the electrons leaving the glass equalled those arriving; the positively charged glass greatly increased the perveance of the tube. For a fuller description the reader may refer to the *General Electric Review*, July, 1920.

Non-uniformity of Temperature.

40. Non-uniformity of the temperature of the filament, to whatever cause it may be due, produces divergences from the simple theory we have so far developed. In tubes with short filaments the end correction for cooling is different at different values of the filament current; increase in the filament current therefore increases the length of the active portion and increases the perveance several per cent. by increasing the cross-section of the space. There is therefore an apparent breach of the rule that the Langmuir part of the characteristic is unchanged by a rise of filament temperature. A secondary effect of this kind may also arise as anode voltage is raised if the anode becomes very hot; for this may raise the temperature of the filament, increase the active length and increase the perveance. The addition to the filament current of a continually growing electron current as the anode voltage is gradually raised may also have secondary effects of this kind.

But non-uniformity of the temperature of the filament, and more especially that due to the cooling of its ends by its leads or supports, has more important consequences than those of the last paragraph. In certain applications of ionic tubes, and particularly in triodes, the nearly straight part of the characteristic formed by the premature bending of the Langmuir part towards the saturation part is advantageous. In some cases the three halves power law holds over only a sixth part of the characteristic. The straightness appears to be due largely to the cooling of the ends of the filament, as will now be explained.

The explanation applies to all forms of tube utilising heated filaments as cathodes and does not depend on whether the filament is straight or folded ; but for simplicity we shall assume that we are dealing with a cylindrical tube possessing a straight axial filament. The middle parts of the filament are surrounded by the anode, which serves incidentally as a good reflector of the heat radiated from the filament and therefore tends to keep up the temperature of the centre portions. The ends of the filament are, on the other hand, strongly cooled by the leads that support it, and also by the opportunity for radiation. Now the electron current that can be drawn from a hot filament increases rapidly with rise of temperature—for example, the saturation current

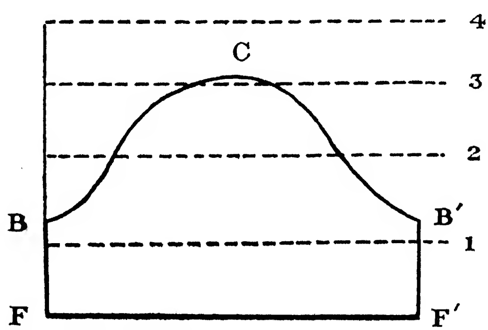


FIG. 239.—Variation of temperature and emission along filament.

per unit length of a certain filament while at a uniform temperature of 2,200° K was 15 mA, and this was attained at 30 V, while the same filament at 2,400° K gave a saturation current of 110 mA at 115 V. A difference of 200° in temperature between the middle and ends of the active part of a filament is not at all unlikely, and therefore in such an instrument the ends yield their maximum current at a voltage of 30 V and cannot contribute more though the voltage be raised. Below that voltage every point of the filament would obey the three halves power law, but above it the ends yield only 15 mA per cm while the centre portions continue to obey the law. But as the voltage is raised further and further, more and more of the filament attains saturation and ceases to contribute further increases of current. Evidently the law which the total current obeys will depend greatly upon the law according to which the temperature is

distributed along the filament, and it is clear that the filament as a whole will not give its final saturation current till the voltage is raised to the saturation value corresponding to the temperature of the hottest part of the filament. All this is represented graphically in Fig. 239. In this Figure, FF′ represents the active length of the filament, and the ordinates of the curve BCB′ represent values of the saturation current per unit length at the corresponding points of the filament.

41. It will be shown below that the fall of potential along the filament is not fundamentally concerned in the present problem, and we shall, therefore, suppose the filament to be at uniform potential throughout its length, and shall represent the voltage between it and the other electrodes by a straight line parallel to FF′. Such lines are drawn and marked 1, 2, 3, 4 in the figure; they are placed at heights that represent to scale the electron current per unit length corresponding to the voltage, that is, their heights are proportional to the three halves power of the voltage. At the position marked 1 the voltage is such that even the parts of the filament opposite the ends of the cylinder are not giving their saturation current, which is equal to FB; and every portion of the filament is yielding current at equal rates per unit length and according to the three halves power law. At the position 2 we have a state of affairs in which large portions near the ends of the filament are giving their various saturation currents, while the middle portion still follows the three halves power law in responding to rise of voltage. At the position 3 almost every part of the filament is giving its saturation current, and at position 4 the current has ceased to respond to changes of the voltage. At any stage the whole current i_a is equal to the area under the line and within the curve.

42. Let l be the active length of the filament, y the length of the centre unsaturated portion in any general position of a line, such as that numbered 2, i_u the electron current being taken per unit length of the unsaturated portion, e_a the voltage throughout the length of the anode, i_a the total current from anode to cathode. Since the filament is all at one potential, we may write

$$i_u = A' e_a^{\frac{3}{2}}.$$

Imagine that the voltage undergoes the increment de_a; the line representing e_a in the diagram will rise and add to the area under-

neath it an amount equal to the length of the unsaturated portion multiplied by the increase in the electron current throughout that portion. The total current increases by the amount di_a, and we have

$$di_a = y di_u.$$

By aid of the preceding equation we obtain

$$\frac{di_a}{de_a} = \frac{3}{2} A' y e_a^{\frac{1}{2}}.$$

This differential equation is the expression of the law connecting i_a, e_a, and y.

In position 1 the value of y is l and we can then integrate the equation and obtain

$$i_a = lA' e_a^{\frac{3}{2}}.$$

In position 2 y varies as the line rises and therefore the equation cannot be integrated till the law of variation is given for the particular filament in use. In position 3 the equation is about to be put out of action by the fact, not expressed in it, that the temperature has a limit. In position 4 the equation is completely out of action; obviously the area of the whole curve FBCB′F′ represents the total saturation current of the tube.

43. The determination of y, which implies the determination of the temperature at every point of the filament, if it had to be done from the fundamental physical laws of the conduction and radiation of heat, would demand a complete knowledge of the dimensions and properties of the filament substance and of the reflecting powers of its surroundings; consequently we leave the connection between i_a, e_a and y in the form of a differential equation. But it is interesting to reverse the process and to enquire how the electron current per unit length (and inferentially the temperature) would have to vary from point to point of the filament in order to give a specified shape to the characteristic. It is conceivable that the designer might control the temperature distribution by, for example, using a filament of variable section, and by altering the configuration of the reflecting surfaces. Therefore let us enquire into the variation of electron current needed to make a large portion of the characteristic curve a straight line. Let the equation of this straight part be

$$i_a = a_0 + a e_a$$

where a_0 and a are constants on this hypothesis.

Then

$$\frac{di_a}{de_a} = a,$$

and therefore the equation

$$\frac{3}{2}A'ye_a^{\frac{1}{2}} = a$$

gives the connection between y and e_a along the straight portion. In terms of the electron current per unit length this becomes

$$\frac{3}{2}A'y(i_u/A')^{\frac{1}{3}} = a.$$

By aid of this equation the curve of Fig. 240 is constructed to exhibit the appropriate saturation current at each point of the filament. The parallel straight line boundaries at the sides of the figure are drawn a distance l apart, where the length l is the

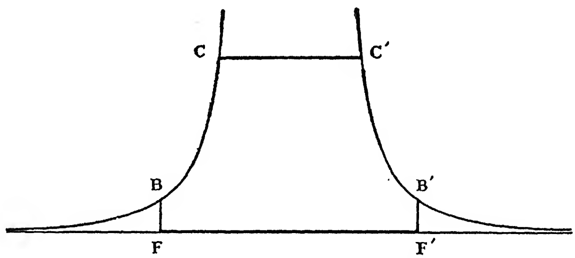


FIG. 240.—Particular case of variation of emission along filament.

active part of the filament. The upper boundary line parallel to the base has yet to be determined. To do this we may return to the differential equation

$$di_a = ydi_u$$

and eliminate i_u by aid of the last equation. First put this in the form

$$i_u = cy^{-3}$$

for brevity and notice that if i_{a1} represent the total current across the vacuum at the instant when the active ends of the filament reach saturation values, then at that instant

$$i_{a1}/l = cl^{-3}$$

in accordance with the above equation. Thus we have for determining c

$$i_{a1}l^2 = c.$$

Now write

$$di_a = yd(cy^{-3})$$
$$= -3cy^{-3}dy.$$

On integrating we get

$$i_a = \frac{3}{2}\frac{c}{y^2} + \text{constant}.$$

When $$y = l,\ i_a = i_{a1},$$

therefore $$i_{a1} = \frac{3}{2}\frac{c}{l^2} + \text{constant}.$$

Hence $$i_a - i_{a1} = \frac{3}{2}c\left(\frac{1}{y^2} - \frac{1}{l^2}\right).$$

Upon eliminating c by aid of the relation found above, we have

$$\frac{i_a}{i_{a1}} - 1 = \frac{3}{2}\left(\frac{l^2}{y^2} - 1\right)$$

or $$\frac{y^2}{l^2} = \frac{3i_{a1}}{2i_a + i_{a1}}.$$

This equation gives the fraction of the active length of filament still unsaturated when the total current is known by measurement. For example, if $i_a = 9i_{a1}$ at the extreme end of the straight portion of the characteristic, we have

$$\frac{y^2}{l^2} = \frac{3}{18 + 1}$$

or $$y = 0{\cdot}397l.$$

Flow of Electron Current in Filament.

44. Uneven heating of a filament arises in large tubes if the external connections are such as to cause the electron current to flow more in one half of the filament than the other. For instance, when the external circuit is connected to the negative end of the filament (as is recommended for the purposes of measurement and record) the electron current will in general be heavier in the negative half of the filament. The difference in large tubes may amount to three or four per cent. on first connection, but clearly this end will soon become hotter and emit more electrons than the other and therefore take a still larger load. A limit is reached after a decided bias has been created. The consequent unevenness of temperature has as its own immediate effect the bending

of the characteristic curve further away from the three halves power curve; and it has the ultimate effect of shortening the life of the filament very considerably.

Effects of the Filament Current.

Potential Drop along Filament.

45. When the cathode is a filament heated electrically there is a fall of potential along the filament amounting to four or six volts in small tubes, and to twenty or forty volts in large ones. It is necessary to examine the effect of this fall of potential upon the characteristic curve. In Fig. 241 FF′ represents the length of filament of a cylindrical diode, F being the negative end. On a perpendicular at F′ set up F′F″ equal to the potential drop along the active part of the filament, say e_f. Draw a perpendicular at F and lines AA′ parallel to FF′ and either intersecting FF′ at a point P or not intersecting it at all. In the former event a triangle APF, shaded in the figure, indicates by its ordinates the voltage between anode and filament in any plane perpendicular to the filament, and evidently relates to cases where the anode potential is lower than that of the positive end of the filament. For the purpose of making a rough approximation we may suppose that motion of the electrons in various plane sections perpendicular to the filament is dictated solely by the voltage shown by the corresponding ordinates of the triangle. Let ξ be the distance from the negative end to any point on the active part of the filament; then the current contributed by an elementary length at the point is

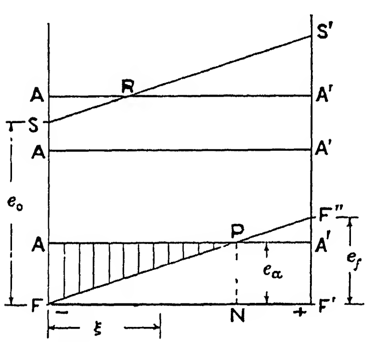


Fig. 241.—Drop of potential along filament.

$$di_a = A'(e_a - e_f\xi/l)^{\frac{3}{2}} d\xi.$$

On integrating from 0 to ξ we obtain

$$i_a = \frac{2}{5}\frac{l}{e_f}A'\{e_a^{\frac{5}{2}} - (e_a - e_f\xi/l)^{\frac{5}{2}}\}.$$

For a point such as P in Fig. 241

$$e_a - e_f\xi/l = 0$$

and therefore

$$i_a = \frac{2}{5}\frac{l}{e_f}A'e_a^{\frac{5}{2}}$$

and this is the solution for the case $e_a < e_f$. In the case when $e_a > e_f$ we apply the integral to the upper line AA′ in the figure, where $\xi = l$. The solution is then

$$i_a = \frac{2}{5}\frac{l}{e_f}A'\{e_a^{\frac{5}{2}} - (e_a - e_f)^{\frac{5}{2}}\}.$$

It is mathematically useful to notice that in all formulæ such as these the factor outside the large brackets is always the same, and the part inside is formed by taking the five halves power of the potential difference at one end of the portion of filament considered and subtracting it from the five halves power of the potential difference at the other end.

46. The essential feature of the solution of the former of these two cases is easily arrived at without integration. For it is clear that the geometrical quantity which we have called the perveance of the tube increases as the active length of the filament increases, and that this length is proportional to e_a while $e_a < e_f$;

therefore $$G \propto e_a$$

and thus $$i_a = Ge_a^{\frac{3}{2}} \propto e_a^{\frac{5}{2}}.$$

As regards the problem relating to $e_a > e_f$, it is evident from inspection that the departure from the three halves power law, which is exhibited in the equation obtained above, must become of smaller significance the higher the anode voltage rises.

47. In Fig. 241, a line SS′ is drawn to represent saturation conditions; S is as far from F as S′ is from F″ in order to indicate that the same voltage is required to draw saturation current from portions of the filament at either end. When the anode voltage line is high enough to cut the saturation line, as at R in the figure, a length ξ_R of the filament contributes at the full emission rate j_1, and therefore yields $\xi_R j_1$ of current. Here, if e_0 represents the saturation voltage indicated by FS in the diagram, we must have

$$j_1 = A'e_0^{\frac{3}{2}}$$

and $$\xi_R = l(e_a - e_0)/e_f.$$

The rest of the filament follows the law expressed by the equation

of the last paragraph, but with $e_0^{\frac{5}{2}}$ substituted for $e_a^{\frac{5}{2}}$. We find for the current from the whole filament the expression

$$i_a = \frac{1}{5}\frac{l}{e_f}A'\left\{e_0^{\frac{3}{2}}(5e_a - 3e_0) - 2(e_a - e_f)^{\frac{5}{2}}\right\}.$$

48. It is perhaps easier to realise the implications of this equation by graphical consideration of Fig. 241 than by inspection of the formula. The ordinates between AA′ and FF′ or between SS′ and FF′, whichever is the smaller, have to be raised to the three halves power and the area of the resulting diagram estimated, in order to watch the growth of the current from one adjustment of anode voltage to another. The immediately evident effect of the gradual saturation of the length of the filament is the absence of abruptness at the saturation knee of the

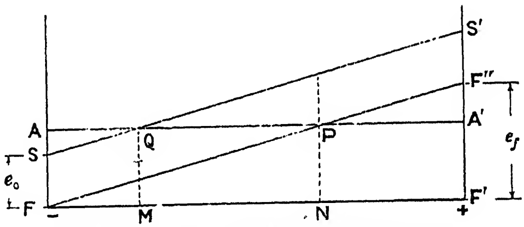


Fig. 242.—Drop of potential along filament.

characteristic curve. It is, however, not to be inferred that this is the sole cause of the slow transition from the sloping part to the level saturation part of the characteristic.

49. There is another important case to consider, namely that indicated graphically in Fig. 242. It is likely to arise in the use of diodes made with long anodes of small diameter, especially if the filament is fine and the temperature high. In such circumstances it is possible to have the P.D. along the filament greater than the voltage that would be needed to produce the saturation current in the same tube if the filament could be all at zero potential. It is then possible to set the anode voltage, as indicated in Fig. 242, so as to draw the saturation current from a length FM, to take current in obedience to the three halves power law from a length MN, and to take nothing save the current due to initial velocities (which will be assumed negligible) from the length NF′. As the anode voltage is raised the line AA′ moves up the page, and while it intersects both the lines FF′ and SS′,

adds at a steady rate to the current collected; for the portion collected from MN remains unaltered, and that collected from FM grows at the rate j_1, in the notation of the preceding paragraph. The yield of the saturation portion is, by the last paragraph,

$$j_1 \xi_M = A' l e_0^{\frac{3}{2}}(e_a - e_0)/e_f$$

which grows at the rate $A' l e_0^{\frac{3}{2}}/e_f$ per unit increase of anode voltage. This, therefore, is the gradient of the $i_a e_a$ characteristic. In order to obtain the expression for the total current taken we have to add to the above that contributed by the portion MN, which is, by the rule stated in § 45,

$$\frac{2}{5} \frac{l}{e_f} A' e_0^{\frac{5}{2}}$$

Hence $$i_a = \frac{1}{5} \frac{l}{e_f} A' e_0^{\frac{3}{2}}(5e_a - 3e_0)$$

when the line AA′ cuts both the saturation and filament voltage lines.

Effect of Initial Velocities.

50. In the above discussion we have neglected the effects of the initial velocities of the electrons. From §§ 18, 19 we know that the effects may be allowed for along the portion FN of the filament by adding a term not directly connected with the varying P.D., and that the current per unit length along the portion NF′ is an exponential function of the P.D. We may therefore write the whole current from NF′ proportional to

$$\int_N^{F'} \epsilon^{-(\xi - l')/b} d\xi$$

which equals $$b \left\{ 1 - \epsilon^{-(l-l')/b} \right\}.$$

Here l' denotes the length of the portion FN. The total current due to the velocities of emission is proportional to the result obtained by adding l' to the above. When this result is written out it is seen that as e_a is increased from zero towards the value e_f the consequent increase in the value of l' contributes to the grand total of current a part that rises somewhat faster than in proportion to the first power of the anode voltage. On the whole, therefore, the effect of the initial velocities is to make the charac-

teristic rise rather less rapidly than in the 2·5 power of the anode voltage when this is less than the potential drop along the filament.

The equations relating to the remainder of the $i_a e_a$ characteristic of the tube under discussion may be written down by combining the results of the preceding paragraphs. From the whole analysis the general conclusion may be drawn that the departures from the three halves power law due to potential drop along the filament are considerable at the lower anode voltages, and whenever saturation conditions appear, but that middle portions of the characteristics are less disturbed unless indeed extreme shapes of tubes or extreme filament conditions are concerned.

Effect of Electric Field of Filament.

51. The perveance of a tube may also be influenced by the filament P.D. in another manner, especially when the filament is folded. Consider the behaviour of any filament when the voltage on the anode is low. The bulk of the electron current comes from the negative end of the filament, as we have seen ; and since the positive end of the filament will produce according to the rules of electrostatics a positive electric field in space near the negative end of the filament, the space charge effect there will be somewhat cancelled, and the perveance of the tube as a whole will tend to be increased. The amount of the increase will vary as the anode voltage changes. The possible occurrence of the reverse result is referred to in § 63. This effect, by the way, must occur even when no voltage is applied to the anode or even in a tube containing no conductor besides the filament. In such a case the potential drop along a straight filament will have appreciable effects, and space charge and filament potential both come into play. On account of this the push on any electron is oblique and not radial, as is shown by compounding the two component forces represented in Fig. 243 at the point P. And if we imagine the electrons moving under these electric fields after emerging from the filament at different places with different speeds, we obtain such curves as are suggested in the sketch. From this it is seen that the vacuum acts as a shunt to the filament in an ordinary incandescent lamp for instance. The heating current carried apparently through the filament goes partly through the vacuum, and this portion may be rendered a quite appreciable fraction of the whole.

Effect of Magnetic Field of Filament Current.

52. There is another way in which the current flowing along the filament may affect the shape of the characteristic curve of a tube. The current produces a magnetic field which when the filament is straight may be represented by circular lines of magnetic force centred on the filament axis. Referring to Fig. 243, the magnetic field at the point P will be directed perpendicularly to the plane of the paper and towards the reader. Consequently the electron at P moving in the direction of the tangential arrow will experience an electrodynamic force tending to urge it in the direction F_2. That is to say, the magnetic field of the heating current helps the space charge in sending electrons back into the filament.

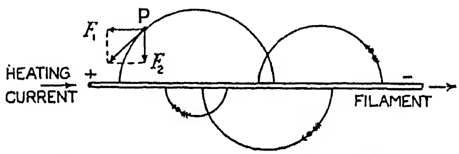


FIG. 243.—Indicating mode of flight of electrons in straight filament lamp.

53. If B represents the magnetic flux at distance x from the axis of a straight filament, and u the velocity of an electron at that place moving parallel to the axis of the filament, the radial mechanical force acting on the electron is Bqu. The electric force F required to balance this is given by

$$Fq = Bqu.$$

But
$$B = 2\mu i_f/x$$

where i_f is the value of the filament current. Therefore $F = 2\mu i_f u/x$ in any units.

Let the velocity u be that corresponding to the average voltage V of emission at a given temperature, then by § 7, we have

$$u = (2Vq/m)^{\frac{1}{2}}.$$

For tungsten at 2,400° K this voltage is 0·31 volt. Hence using electromagnetic units we have

$$u = (2 \times 0{\cdot}31 \times 10^8 \times 1{\cdot}77 \times 10^7)$$
$$= 3{\cdot}32 \times 10^7 \text{ cm/sec.}$$

Let $x = 0{\cdot}0037$ cm be the radius of a filament carrying an ampere, then in electromagnetic units

$$F = \frac{2 \times 0{\cdot}1 \times 3{\cdot}32 \times 10^7}{0{\cdot}0037}$$
$$= 1{\cdot}80 \times 10^9.$$

That is to say a potential gradient of eighteen volts per centimetre is required to pull an electron out of the magnetic field near a filament when the electron is moving with the average kinetic energy parallel to the filament from the low potential end to the high, and when the filament is of the dimensions and temperature indicated.

54. Since we know from § 13 that the quantity $i_f/(2x)^{\frac{3}{2}}$ is a function of the temperature only and is equal to $\sqrt{(EJ/R)}$, we may eliminate the filament current from the equation for F and obtain

$$F = 4\mu u \sqrt{(2x\, EJ/R)}.$$

Thus the potential gradient required increases as the square root

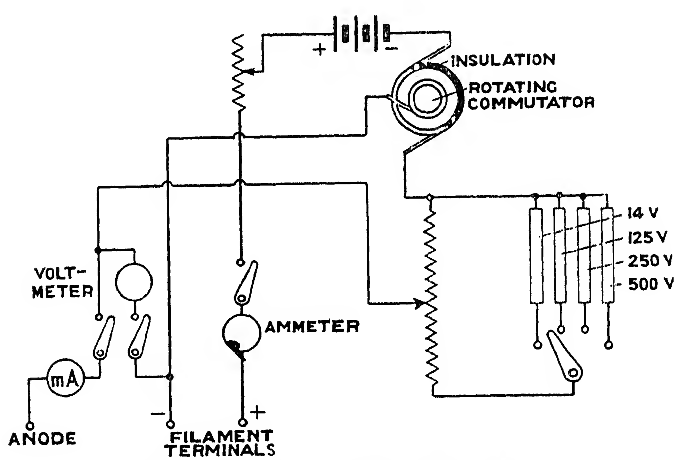


Fig. 244.—Elimination of effects of filament current.

of the diameter of the filament in the extreme case when the electron is moving parallel to the surface of the filament very shortly after emission.

Measurement of the Space Charge Effects.

55. Measurements made in highly evacuated tubes have fully vindicated the space charge theory. A long series is described by S. Dushman in the *Physical Review*, 4, 121, 1914, and another series by Schottky in the *Physikalische Zeitschrift*, 15, 624, 1914. In such measurements it is advantageous to eliminate the effects of the heating current in the filament by interrupting this circuit

and connecting it to the measuring apparatus many times a second by aid of a rotating commutator or equivalent method. These methods were first used by O. von Baeyer and subsequently by Schottky. Fig. 244 is the diagram of the circuit used in the General Electric Laboratory. All the effects due to the P.D. along the filament, the magnetic field of the filament current, and the space effect of the electric field of one part of the filament on another, are eliminated by this mode of measurement.

The Saturation Current.

56. The fundamental space charge theory developed in the preceding pages holds good so long as there is a superfluity of electrons emitted at each point of the cathode; but when the voltage applied between anode and cathode becomes great enough to take from any part of the cathode all the electrons being emitted there, the current from the tube as a whole is not strictly proportional to the three halves power of the voltage. If the voltage be gradually increased there is a transition period, as we have seen, during which all parts of the cathode in turn attain the saturation condition. It is to be expected that when all parts of the cathode are having drawn from them their complete emission the current will not increase however high the voltage be pushed, the vacuum being supposed perfect. The characteristic curve will then become level. But as a matter of experience it is very difficult to obtain a level curve; it usually rises as the anode voltage increases, though at a gradually diminishing rate. There are several known causes for this phenomenon.

57. One cause is the gradual rise of the temperature of the cathode due to radiation of heat from the anode. When the anode voltage is raised a step the electron velocity is increased and the greater bombardment brings the anode to a higher equilibrium temperature. If the increase in the heat radiated to the cathode is enough to raise its temperature appreciably the emission increases and therefore the saturation curve rises. According to Langmuir, this effect is apt to be important with low temperature cathodes, such as the lime-coated cathode of Wehnelt, for the reason that the received heat radiation is reradiated less by a low temperature surface than by a high temperature one.

58. Another consideration, which may enter especially at very high voltages, arises from the possibility pointed out by Schottky of increasing the emission from a metal by applying at its surface an electric field approaching the magnitude of a million volts per centimetre.

59. The most commonly occurring causes of departure from the ideal saturation condition are heterogeneity and irregularity of the cathode surface. Langmuir describes an experiment in which a thoriated filament in a very high vacuum was heated at 1,900° K until its surface became completely covered by thorium diffused from the interior. When this was heated at a rather low temperature it gave a fairly level saturation curve over a wide range of voltages. The filament was then heated for a few minutes to 2,900° K so that all the thorium distilled from the filament and left it with a pure tungsten surface. The emission was now very much smaller than it was at the same temperature with the thorium surface but the saturation curve was very definite. On the other hand, when the surface, by appropriate heat treatment, was left partly covered with thorium and partly with tungsten no well-defined saturation current could be obtained; the current rose continually as the voltage was raised. It is probably for such reasons as these that Wehnelt cathodes rarely give a well-defined saturation current. And to this must be added the effects of geometrical irregularity of the surface of a lime-coated filament.

60. Even when a tungsten filament has only slight traces of thoria upon it very high external fields are required to get saturation. This effect is at least in part associated with contact potential phenomena. A speck of thoria on a tungsten surface becomes and remains positive relative to the tungsten because it emits electrons faster, or in other words, the tungsten becomes negative relative to the thoria. The electric field very near to and outside the thoria is thus determined by the tungsten, and is held relatively negative in spite of the application of large anode potentials.

Other Anomalies in Diodes.

61. Leaving aside for a moment the effects of minute traces of gas, which offer many and difficult problems, we may put under this heading a class of abnormality that has already been referred

to occasionally and is well illustrated by the following experiment of Langmuir's. In this experiment an emf of a few hundred volts was applied between anode and cathode and the filament temperature gradually raised or, alternatively, the filament was first glowed and the anode voltage gradually raised. The tube operated in the normal manner but the current was limited to a rather small value. If now a switch in the anode circuit was opened for an instant and closed again it was found that the whole character of the discharge had changed. The current to the anode was much greater, but the heat developed at the anode was less than before, and considerable heat was liberated on the glass walls of the bulb. The explanation of these events is as follows : In the first part of the experiment the glass walls of the tube became gradually covered with a deposit of slowly-moving electrons, and this negative charge, once settled, prevented other electrons striking the glass and tended to neutralise at the cathode the field produced by the anode ; therefore the current was of low value. When the anode current was stopped and restarted suddenly high velocity electrons were thrown against the glass and produced a secondary emission of low velocity electrons greater in amount than the striking electrons ; hence the glass became charged positively and attracted more electrons. This cumulative process went on until the glass was at a potential nearer that of the anode than of the cathode and the number of electrons leaving the glass was equal to the number striking it. Besides attracting a continual stream of high-speed electrons to itself the charged glass increased the perveance of the tube by means of the field it exerted at the cathode. As regards the thermal effects, the slow-moving electrons received by the anode in the second part of the experiment generated less heat than the quickly moving ones received in equal number by the glass.

62. The preliminary stage in this experiment illustrates what usually happens in any diode. The glass, or even a conductor if insulated, tends to collect electrons until it has acquired a potential sufficiently below that of the cathode to prevent more alighting upon it. This potential is usually of magnitude about two volts. Thereafter its field assists that of the space charge in repelling electrons into the cathode, that is to say, it reduces the current-carrying capacity of the tube. It is further to be noticed that in such cases the field at the cathode may be so negative in

spite of the application of a large anode voltage that the effect of the initial velocities of the electrons may come into prominence and marked deviations from the three halves power law arise. In fact it is easy to make tubes with well separated small electrodes in which the perveance becomes zero when the glass becomes charged.

63. The effect of the potential of one part of the filament on the space charge near another part has already been alluded to in § 51, but it may be mentioned again under this heading. As an instance consider a tube with a V-shaped filament placed between two plane electrodes connected together to form the anode. If the anode circuit be joined externally to the negative end of the filament the positive leg of the filament produces along the other leg an electric field that tends to improve the perveance of the space. But if the anode be connected to the positive leg of the filament the negative leg tends to diminish the current-carrying capacity. Suppose that the anode voltage be 100 and the filament voltage 6, then in the former case the mean potential difference between the electrodes is 97, and in the latter 103—yet the anode current may be smaller in the latter case. All such cases and the similar ones arising out of the charging of the glass walls of bulbs are instances of what might be called the "third body effect" which will be discussed more fully in a later paragraph.

Algebra of Diode Characteristics.

64. From the preceding pages we may draw the conclusion that ideal obedience of the volt-ampere characteristic curve to the three halves power law, even in the case of an absolute vacuum, can be expected only if the cathode is perfectly uniform in form and material and is all at the same temperature and the same electric potential. The main causes of departure have been described, and it is clear that many of these causes may interact with each other in ways too various and complicated to be followed out in detail. Moreover, in practical tubes there is always some gas present, and the effects of this remain to be discussed in later paragraphs. Fortunately many of the highly evacuated tubes occurring in practice possess characteristics which admit of simple and yet sufficiently approximate algebraic treatment in the circumstances in which the tubes are usually used.

We shall now develop a few approximate equations with a view to providing general means of solving the more elementary problems that arise when ionic tubes are introduced into electrical circuits.

65. We have seen that the electrons moving from the filament to the anode across the vacuum and constituting the convection current give rise to an electric field tending to cause the electrons near the filament to return to it. Let us represent the effect of this restraining field as a back emf opposing the main battery in the anode circuit, and somewhat analogous to the back emf of polarisation occurring in a liquid electrolyte. We have seen that the density of the space charge is greatest near the cathode, and we therefore conclude that the back emf may be conceived to be localised near the filament. We may picture it in a sketch by drawing the conventional symbol for a voltaic cell at such a position as e in Fig. 245. The magnitude of this back emf will be represented by e_s. The voltage between anode and cathode will be represented by e_a, the current through the diode by i_a.

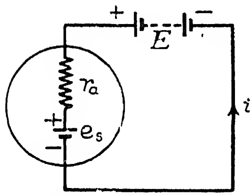


FIG. 245.—Representation of diode circuit.

As the voltage is varied, the current is found to change somewhat in the manner indicated by the characteristic curve of Fig. 246. The portion OK_1 obeys approximately the theoretical space charge equation

$$i_a = Ge_a^{\frac{3}{2}};$$

where G is constant if the filament voltage is small, as explained in § 34. The portion

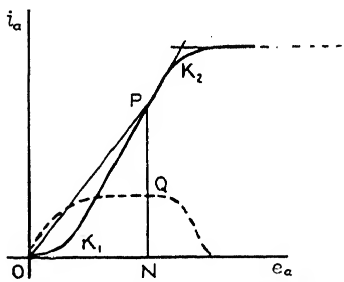


FIG. 246.—Characteristic and gradient curves.

K_1K_2 departs from this equation because there are inequalities of temperature existing between the ends and the central portions of the filament which lead to the various parts of the filament reaching the saturation condition in turn; the portion beyond K_2 and parallel to the voltage axis is unvarying because all the electrons emitted from every part of the filament are being

carried to the anode and no more are available. The points K_1 and K_2 are called the ankle and the knee of the characteristic curve.

66. If the nearly straight part be replaced by a straight line, as in Fig. 247, and this equivalent line produced backwards to cut the voltage axis at T, its equation may be written

$$i_a = a_a(e_a - e_s)$$

where e_s is the value of OT and a_a is the gradient of the line, that is, the tangent of the angle of slope. The quantity a_a has the dimensions of a conductance and is constant; the quantity e_s is also a constant, and may be interpreted as the back emf. The tube may therefore be said to obey Ohm's law formally. When convenient we shall write r_a for the reciprocal of a_a and shall call r_a the resistance of the tube.

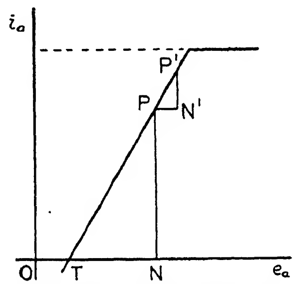


FIG. 247.—Idealised characteristic.

67. Turning now to the portion of the curve corresponding to voltages too low to draw the saturation current from even the coldest parts of the filament, taking any point P and drawing the tangent line as shown in Fig. 248, we find from the equation

$$i_a = Ge_a^{\frac{3}{2}}$$

that the gradient of the tangent line is, if G is constant,

$$\frac{di_a}{de_a} = \frac{3}{2}Ge_a^{\frac{1}{2}}$$

We shall, for short, call this a_a.

Then $a_a = \frac{3}{2}i_a/e_a$.

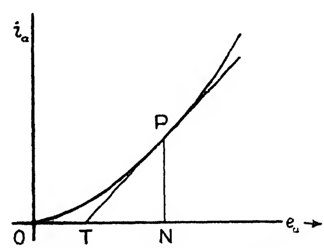


FIG. 248.—Langmuir portion of characteristic.

We have, also, OT = ON — TN, if PN be the ordinate of the point P. Call the intercept e_s,

then $$e_s = e_a - i_a/a_a$$

or $$i_a = a_a(e_a - e_s).$$

Also $$e_s = e_a - \frac{2}{3}e_a = \frac{1}{3}e_a.$$

The equation for i_a becomes the equation of the curve when a_a

and e_s are given their values in terms of the co-ordinates of the point P, but it becomes the equation of the tangent line if a_a and e_s be held constant. In many physical problems we may identify a very small length of the curve with an element of the tangent at a point within that length, and may therefore regard a_a as instantaneously the conductance at the point P and e_s as the back emf; in such cases a_a may be called the differential conductance of the tube at the adjustment in question and the diode may be treated in calculations just as if its perveance were made up of a constant back emf e_s and a conductance. The cases in which this simplification cannot be allowed are those in which the curvature of the characteristic is the property, or one of the properties, being utilised in the experiment under investigation.

68. In order to remove any possibility of misunderstanding in the use of the phrase "conductance of a tube" Fig. 246 may be referred to. Here lines OP and PN are seen containing the angle PON. The tangent of this angle is the ratio i_a/e_a and this has by some writers been taken as the conductance of the tube. From this point of view the conductance of a diode is variable with the applied emf. On the other hand the internal or differential conductance at the point P in the sense used in this book is the gradient of the curve at the point P, and is obtained by drawing the tangent line at P as already explained. If this conductance at all such points as P be represented by a length NQ on the ordinate of P, the dotted curve is obtained.

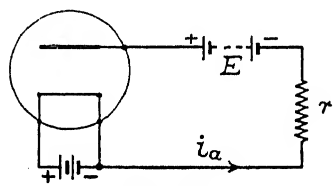


FIG. 249.

Diode in Series with Resistance.

69. In Fig. 249 let a battery E, a resistance r and a diode of resistance r_a and back emf e_s be connected in series. In order to calculate the current that flows (always supposing that the saturation current is not being drawn) we add together the resistance in the circuit and also the emfs and divide the latter by the former. Thus in Fig. 249

$$i_a = \frac{E - e_s}{r + r_a}.$$

This equation may be written

$$r_a i_a + e_s = E - r i_a.$$

Clearly each side of this equation is equal to the potential difference e_a between the electrodes. This example shows plainly that two constants are required to define the chief properties of a diode. But on the other hand it should be noticed that if we are not going to apply the above equations to the curved part of the characteristic and if we have to calculate the change in current produced by a change in applied voltage confined to the straight part of the characteristic we write

$$\delta i_a = \frac{\delta E}{r + r_a}$$

and in this case need to know only one parameter of the tube, namely, its resistance.

Numerical Examples.

70. Let the point P in Fig. 247 represent the state of affairs in a given tube. When the P.D. between its terminals is 80 V and the current through it 4 mA let the differential resistance be 12,500 Ω. Let it be required to calculate the alteration of current produced by an alteration of 10 V in the terminal P.D. We have

$$\begin{aligned} \text{current difference} &= \frac{\text{voltage difference}}{12{,}500} \\ &= 10/12{,}500 \\ &= 0{\cdot}8 \text{ mA.} \end{aligned}$$

It will be noticed that the apparent resistance offered by the tube to the current of 4 mA is $80 \div 4 \times 10^{-3}$ which equals 20,000 Ω, while, if the current is increased by $0{\cdot}8 \times 10^{-3}$ A this resistance is $90 \div 4{\cdot}8 \times 10^{-3}$ which equals 18,750 Ω. Evidently the apparent resistance of the tube changes considerably though the differential resistance is constant.

Let now a resistance of 2,500 Ω be connected in series with the tube. A current of 4 mA will flow if the battery voltage be made 90; for this current flowing in 2,500 Ω causes a potential drop of 10 V, which leaves 80 V across the terminals of the tube. Let it be required to find the effect of changing the battery voltage by 20 V, supposing the straight part of the characteristic

not departed from. According to the differential equation above we shall have

$$\text{current difference} = \frac{\text{voltage difference}}{2{,}500 + 12{,}500}$$
$$= \frac{20}{15{,}000}$$
$$= \frac{4}{3} \text{ mA.}$$

The difference may be positive or negative. It will be noticed that by working with the differences of current and voltage the necessity of knowing the value of the back emf is avoided.

71. As another example, suppose a resistance $r = 5{,}000\ \Omega$ connected in series with this tube, and E made equal to 60 V. What current will flow assuming saturation not reached ? We can use the equation

$$i_a = \frac{60 - e_s}{5{,}000 + 12{,}500}$$

which involves e_s. In order to find e_s we utilise the original data that 80 V gives 4×10^{-3} A through the tube and that its differential resistance is 12,500 Ω. Referring to Fig. 248 we have

$$\text{TN} = \text{PN}/a_a$$
$$= 4 \times 10^{-3} \times 12{,}500$$
$$= 50 \text{ V.}$$

Hence $\quad \text{OT} = e_s = 30 \text{ V.}$

The question above can now be answered. We have

$$i_a = \frac{60 - 30}{5{,}000 + 12{,}500}$$
$$= \frac{3}{1{,}750}$$
$$= \frac{12}{7} \text{ mA.}$$

Properties of Diodes containing Gas.

Some Data of the Kinetic Theory.

72. In high vacuum work it is becoming customary to record the pressures in dynes per square centimetre. A pressure of one dyne per square centimetre is called a bar and is about equal

to a millionth of a standard atmosphere, being, in fact, very nearly a millionth of the pressure due to 750 mm of mercury at 0° C. in latitude 45 degrees at sea level. Pressures are also often recorded in thousandths of a millimetre of mercury which are called microns and given the symbol μm. Thus one bar is equal to 0·75 micron. A vacuum is considered to be "high" when the pressure is between 0·001 and 1 micron; the highest attainable vacuum at the present day is of magnitude about 0·0001 micron, that is about one-tenth of a millimicron or about one-tenth of a millibar. In an ideal gas there are $2{\cdot}67 \times 10^{19}$ molecules per cubic centimetre at the standard atmospheric pressure. Hence, in the vacuum of one-tenth of a millibar there are 2,670,000,000 molecules per cubic centimetre at 0° C. The mass of one cubic centimetre of hydrogen under the conditions stated is $0{\cdot}83 \times 10^{-14}$ gram.

According to the kinetic theory of gases the molecules of a gas at ordinary temperature are moving about at great speeds in random directions and with frequent collisions. The mean velocity is proportional to the square root of the absolute temperature. In hydrogen at 0° C. it is $1{\cdot}84 \times 10^5$ centimetres per second, in argon it is $0{\cdot}413 \times 10^5$ centimetres per second. The number of collisions per second at ordinary pressure is, in hydrogen, about 10^{10} and in argon about 4×10^9. The mean distance travelled between collisions is called the "mean free path." It varies inversely as the pressure. In a paper on the kinetic theory in the *General Electric Review*, p. 1042, 1915, S. Dushman gives a large number of valuable kinetic data; from these it appears that for most gases the mean free path under standard conditions is about 10^{-5} cm, but for hydrogen it is 1·7, and for helium 2·7 times that distance. For most gases the mean free path at a pressure of 1 bar is about 10 cm, for hydrogen it is 17 cm, for helium 27 cm; in a tungsten lamp, where the pressure may become about a hundredth of a bar, the mean free path is several metres.

Electrons in Gases.

73. When free electrons are introduced among the molecules of a gas at ordinary pressure they quickly combine with the neutral molecules and form negative ions which can be carried through the gas by the application of an electric field. This formation of negative ions of molecular size is the most likely

interaction if the electric field is less than 10 volts. But if a sufficiently strong field is already present the electrons as they enter the gas may be dashed against molecules with sufficient force to knock an electron out of the neutral molecule. This process is called ionisation by collision, and results in the production from a single molecule of another free electron (or electrons) and a positive residue called a positive ion. The strong tendency of these positive and negative ions to recombine can be resisted only by the application of a sufficiently intense field; and evidently such a field may give the new electrons momentum to cause these to ionise still other molecules by collision. That is to say an intense field may lead to the catastrophe of a glow discharge, or even a spark discharge through the gas. The subject is a large one and for information the reader should refer to text-books on conduction of electricity through gases. The main property that concerns our immediate study is the fact that the electron speed needed to ionise gases by collision is different for each gas. When measured by the potential difference through which an electron must rise in order to gain the necessary velocity it is called the ionising potential of the gas. The ionising potential of nitrogen is 18 volts, of argon 17 volts, of carbon monoxide 14 volts, of mercury vapour 10·4 volts, of helium 21 volts. Davis and Goucher have shown that hydrogen ionises in two stages, one at the potential of 11 volts and the other at 15·8 volts. For most of the gases occurring in ionic tubes the ionisation phenomena begin to be important when the voltage applied between the anode and cathode rises to between 12 and 25 volts. At voltages higher than the ionising potential the amount of ionisation increases rapidly with voltage, and being directly connected with the mean free path and the collision frequency it is approximately proportional to the pressure of the gas and to the density of the electron current. It is also proportional to the distance between the electrodes, and depends somewhat on the rate of recombination of positive ions and electrons.

Convection Currents due to Ions.

74. The word "ion" means "wanderer"; in a general sense it is used to convey the idea of a molecule charged negatively or positively, or of a free electron, or of a group of molecules bearing a charge. In ionic tubes we have to deal with the light ions

occurring in high vacua which are always single molecules or electrons.

In an electric field an ion gathers speed and if the gas is rarefied travels on the average several centimetres before it is interrupted. Let us compare the motion which a parallel electric field gives to ions possessing equal charges q but having different masses m, m', etc. Their motion constitutes a convection current of which a part of density i is due to the ion m, a part i' to the ion m', and so on. Let $n, n', \ldots$ be the number of ions per cubic centimetre of the several kinds, v the potential difference between two imaginary planes perpendicular to the field, $u, u', \ldots$ the velocities. Then the energy equations, which state that the kinetic energy gained is equal to the electrical work done, are for two kinds of ions

$$\frac{1}{2}mu^2 = qv$$

$$\frac{1}{2}m'u'^2 = qv$$

and the current density equations are

$$i = qnu$$
$$i' = qn'u'.$$

We obtain

$$i = qn\,(2vq/m)^{\frac{1}{2}}$$
$$i' = qn'\,(2vq/m')^{\frac{1}{2}}.$$

Division gives

$$\frac{i'}{i} = \frac{n'}{n}\sqrt{\frac{m}{m'}}.$$

We proceed to interpret these equations, and in order to assist the explanation we shall apply them to a space containing electrons and mercury molecules moving between two parallel electrodes.

75. The mass m' of the mercury ion may be taken as 360,000 times the mass m of the electron. The energy equations therefore prove that the velocity gained by the electron in moving through a given potential difference is 600 times that gained by a mercury ion moving through the same potential difference. The convection current due to each can be calculated from the current equations if we know the ratio of the numbers of the ions per cubic centimetre. If, for instance, the ion densities are equal, that is if $n = n'$, the mercury convection current is only 1/600th

part of the electron current. If, on the other hand, the currents are equal there must be 600 times as many mercury ions as there are electrons. The last of the equations in § 74 summarises the results compactly. To put the matter in words: Slowly moving ions convey across a square centimetre less electricity per second than an equally dense distribution of quickly moving ions if each bears an equal charge, and the ratio of the currents is the inverse ratio of the square roots of the masses.

76. These results will have application later in ionic tubes in which positive gaseous ions appear in company with electrons. They show immediately that positive mercury ions sufficient to neutralise the negative space charge at any point of an ionic tube will, on account of smallness of the velocities they gather in the same electric field, convey an inappreciable current compared with that carried by the quicker electrons. As for positive hydrogen ions, since their mass is 3,600 times that of the electrons, the current they will convey while completely neutralising the space charge in any region will be one-sixtieth part of that conveyed by the electrons.

77. On the other hand, if the current carried by the positive ions is about equal to that carried by the electrons the gas pressure must be such that on the average each electron shall have one collision with a molecule on passing from cathode to anode—voltages well above the ionising potential being assumed. That is to say the mean free path of the electrons in the gas must be of the same order of magnitude as the distance between the electrodes. Thus, if in a certain tube containing hydrogen the electrodes are a centimetre apart the gas pressure should be about 100 bars.

Bombardment by Ions.

78. The blow delivered to an electrode or to the glass of a tube by a molecule or ion is equal to the momentum of the molecule or ion at the instant of striking. The aggregate of the blows constitutes the process called bombardment. If we consider two kinds of ion, such as hydrogen and mercury, since they carry equal charges they gain the same kinetic energy in travelling through the same electric field. Let their masses be m, m' and their velocities u, u'. Then

$$\frac{1}{2}\,mu^2 = \frac{1}{2}\,m'u'^2.$$

The ratio of the blows they respectively deliver to the cathode is the ratio of their momenta and is

$$\frac{mu}{m'u'}=\sqrt{\left(\frac{m}{m'}\cdot\frac{mu^2}{m'u^2}\right)}=\sqrt{\frac{m}{m'}}.$$

That is to say, the impulse delivered by a molecule is proportional to the square root of its mass. It follows that the heavier molecule produces the greater disintegration. The three inert gases helium, argon and mercury have all been used for making tubes "soft"; the disintegration of the filament is least with helium and greatest with mercury.

Effects of Gas on the Operation of Diodes.

79. We have seen that in perfectly evacuated tubes designed to pass convection currents of magnitude one milliampere or greater the current and the voltage applied between the electrodes are related in three several ways according to the magnitude of the voltage. This is especially clear when the electrodes are ideally uniform and so placed as to have little external field, and when the potential drop along the filament is negligible; for then the characteristic curve possesses three fairly well defined sections linked by transition curves. In the small tubes of commerce the three sections comprise: first, one correlating the voltage and the current flowing when the emf applied to the anode is not more than a few tenths of a volt positive, or when it is negative. The relation between current and voltage in this section is governed by the distribution of initial velocities. Second, a section relating to all positive voltages between about one to two volts positive, and the voltage at which the saturation stage begins. The current in this section is governed by the mutual repulsion of the electrons constituting it and is proportional to the three halves power of the voltage. It may be called the Langmuir section. The third section is ideally a level portion exhibiting the saturation condition. The value of the current here is governed solely by the emission from, and therefore by the temperature of, the cathode.

We have now to examine how these three sections of the characteristic are affected by the presence of gas in the tube. We can apply only broad considerations since in all cases the result depends upon the configuration of the electrodes, on the

exact pressure and nature of the gas, and, to some extent, on the nature of the surface of the electrodes.

Currents at Negative Anode Voltages.

80. Electrons emitted into a rare gas at velocities corresponding to ordinary filament temperatures are hampered in their passage from cathode to anode for at least two reasons. In the first place collisions may occur between the electrons and the molecules, that is to say, the mean free path of the electrons is reduced and in consequence the journey of the average electron takes more time than in a perfect vacuum. In the second place some of the electrons striking the molecules become attached and form relatively heavy negative ions. Since these move with the relatively slow speed of the molecules the average time taken by these negative charges in transit from cathode to anode is greatly prolonged. This implies that the current will be reduced and also that the space charge effect, such as it is at these small current densities, will be increased. Indeed, by the reasoning of § 75 we are shown that a given current of slowly moving ions must have greater ionic density than the same current of quickly moving ions, and therefore that the former must have the greater space charge effect. On the whole the presence of gas tends to reduce the current as compared with that through a vacuum.

The Currents at Moderate Voltages.

81. The medium section of the characteristic extends from one or two volts positive to a voltage near that at which saturation is attained, and in general the ionising potential of the gas falls between those limits. It is, therefore, convenient to divide the Langmuir section of the characteristic into two portions, one of which ranges up to the ionising potential. In this region none of the electrons gain sufficient speed to ionise the molecules, and the only consequences of such collisions as occur are the lengthening of the time of transit of the average electron and the formation of negative ions of molecular speed.

The reasoning applied to the section dominated by the initial velocities (§ 80) is again appropriate and we conclude that the presence of gas will reduce the current flowing. Observation

shows that at, say, 5 volts, the reduction may amount to nearly 40 per cent. The effect is more pronounced at greater pressures and becomes prominent when the mean free path of the electrons in the gas is of about the same value as the distance between anode and cathode. In mercury vapour this occurs at a pressure of about 60 microns in a small tube with electrodes separated about one centimetre.

Ionisation by Collision.

82. The second portion of the Langmuir section is entered upon immediately ionisation begins. If the anode voltage is very little above ionising potential an electron moving from the cathode will not gain the ionising speed until near the anode. Therefore positive ions will be formed near the anode and begin to move towards the cathode under the electric field. If on the way a positive ion meets a slow moving electron they may combine to form a neutral molecule, but while un-neutralised it plays a part in cancelling the field due to the negative space charge. A stage is soon reached in which a certain number of the ions are created per second by collisions near the anode and gradually neutralised at the same rate on their way to the cathode. Thus a definite electron current greater than otherwise would have been obtained flows at the given voltage. On applying a somewhat higher voltage the electrons acquire the ionising speed well before reaching the anode and therefore more collisions occur. In addition the positive ions moving towards the cathode get nearer to it and, besides, gain high speeds and are less easily neutralised. That is to say, the volume near the anode in which ionisation takes place is continually growing with the increase of the applied voltage while the volume in which recombination can take place is gradually decreasing. In consequence, the concentration of the positive ions in the space between the electrodes increases as the voltage is raised, and the field of the negative space charge is correspondingly cancelled. At a still higher voltage account has to be taken of the fact that the electrons knocked out of the molecules in the process of ionisation by collision may themselves acquire ionising velocity and so add to the growth of the concentration of positive ions. Clearly the amount of ionisation must grow with the voltage at an ever-increasing rate.

Departure from the Langmuir Curve.

83. Returning to an anode voltage near the ionising potential where, as explained in § 81, the current is less than would be given by the three halves power law, it is to be expected that after the beginning of ionisation the current will overtake the three halves power of the voltage and as the voltage is raised will surpass the value obtained in vacuum. In other words, the characteristic yielded by a tube containing residual gas will cross from below to above the Langmuir curve when the voltage is a little higher than ionising potential. After that it will leave the Langmuir curve behind more or less rapidly according to the amount of gas

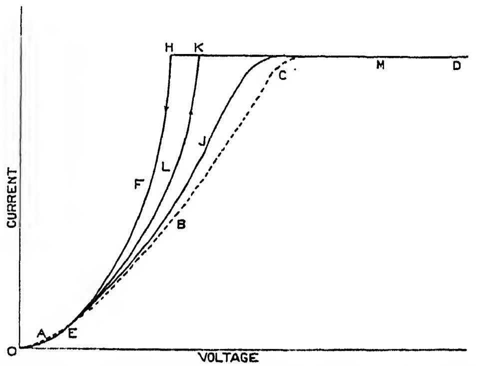


FIG. 250.—Characteristic curves obtained when gas is admitted in two stages.

present. Fig. 250 gives a free-hand sketch of curves belonging to a tube which begins in a perfectly evacuated condition and has gas admitted to it in two stages. The dotted curve OABC has ordinates proportional to the three halves power of the abscissæ. The ionising potential occurs at the point E. When a small amount of gas is admitted the curve OEJ is obtained; and a curve like ELK is obtained when, say, more than a micron of mercury vapour is admitted. Such a curve may become exceedingly steep, and full saturation may be reached by a sudden jump when the voltage is increased by a small fraction of a volt beyond that indicated by L. In such cases it is found as a matter of experiment that when the voltage is reduced the representative point of the characteristic travels from K to some point H before any reduction of the current below the saturation value occurs;

and then further reduction produces a rapid fall of current to a point F on a different curve from that obtained with ascending voltages. It is on account of this property that practical difficulties arise in the use of ionic tubes at sensitive settings.

84. We may compute the influence of a trace of gas at voltages corresponding to the point J in Fig. 250, which is well above the ionising voltage, in the following manner. Let us suppose that the distance between cathode and anode is about 1 cm and that the pressure of mercury vapour is about 0·01 micron, then the mean free path of an electron in the gas is 6,000 cm. Thus, in travelling 1 cm between the electrodes only one six-thousandth of the electrons may be expected to experience collision. We may suppose that at each collision a positive ion is produced. The velocity of mercury ions is one six-hundredth that of electrons in the same field; therefore the number of positive ions existing at any instant when the state is steady is 600 ÷ 6,000, or one-tenth of the number of electrons. This result implies that one-tenth of the negative space charge would be neutralised and that the anode current would be approximately 10 per cent. larger than under the same voltage in a perfect vacuum. The computation proves that the presence of a very minute quantity of gas should be easily perceptible by electrical means; in fact this method of testing for the presence of ionisation is stated by Langmuir, to whom the above computation is due, to be the most sensitive and most practical method of detecting ionisation. It is recommended for determining if a given tube will give steady reproducible characteristics and have a long life.

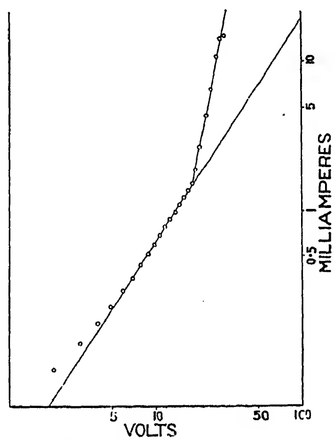


FIG. 251.—Effect of gas on characteristic.

85. The greater the pressure of gas the more pronounced is the departure from the Langmuir curve, and it begins more sharply at relatively high pressures than at low. Ionising potentials may be measured accurately using the relatively high pressures of

50 or 100 microns and duly allowing for the effects of initial velocities. But keeping to the much smaller pressures usual in ionic tubes it is found that after ionisation has fully begun the current, in its departure from the three halves power law, follows approximately another power law, the difference between this new index and the 3/2 of the Langmuir curve being a measure of the pressure of the gas. It may be explained here that the best easy way of determining the numerical value of the index from a given set of observations is to plot the logarithms of the voltage readings as abscissæ and the logarithms of the current readings as ordinates. Then the index sought is the gradient of the line drawn in a mean position among the plotted points in the

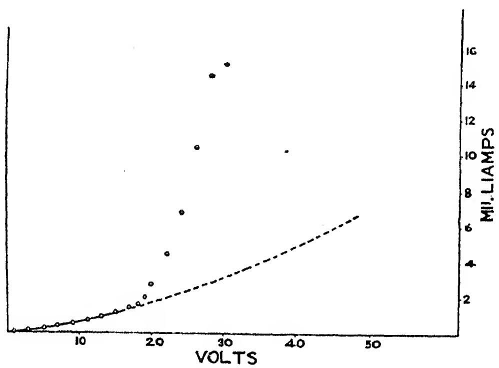


FIG. 252.—Effect of gas on characteristic.

region under consideration. As an illustration of the method consider the results recorded on the curve of Fig. 251, which is the logarithmic plot from an experiment of Langmuir's. The gas in the tube was largely carbon monoxide evolved from a nickel anode. A straight line with a gradient of 3/2 has been drawn through the points that lie just below the ionising voltage. The departure of the current from the three halves power of the voltage is then plainly exhibited. It is to be observed that ionisation begins at between 15 and 20 volts, and becomes prominent about 18 volts. At higher voltages the observed points lie approximately on a steep straight line having a gradient of 5·1. The ordinary characteristic of this tube is given in Fig. 252.

The Gas Line.

86. Let i', e' represent the current and voltage at any point of this gas line and i_0, e_0 the current and voltage at the point of intersection of the gas line and the vacuum line; then we may write the equation of the gas line in the form

$$\log i' - \log i_0 = \alpha (\log e' - \log e_0)$$

or

$$i'/i_0 = (e'/e_0)^\alpha,$$

where α is almost independent of current and voltage but has application only above the ionising point. For approximate purposes we may identify e_0 with the ionising potential and then from the definition of the perveance G of the tube (§ 33) we have

$$i_0 = Ge_0^{\frac{3}{2}}.$$

The quantity G is a purely geometrical quantity in the ideal perfectly evacuated tube and is almost constant in a tube containing only traces of gas. If we apply the same definition to values above the ionising point we shall obtain a spurious or apparent perveance G' given by

$$i' = G'e'^{\frac{3}{2}}$$

Hence

$$G'/G = (e'/e_0)^{\alpha - \frac{3}{2}}.$$

Thus the apparent perveance when ionisation is taking place varies with the voltage applied. The index $\alpha - \frac{3}{2}$ is found to be approximately proportional to the pressure of the gas. When the vacuum is practically perfect $\alpha = \frac{3}{2}$; when there is sufficient ionisation to cancel the field of the negative space charge α is very large.

87. Since the ionising potential of mercury vapour is 10·4 volts a measurement of current and voltage at about 10 or 11 volts enables the geometrical perveance of the tube to be determined even when gas is present. It should also be noticed that measurements of current and voltage at two distinct points above the ionising point and below the saturation point lead to the determination of the index of gas pressure $\alpha - \frac{3}{2}$ without an exact knowledge of the ionising potential or the geometrical perveance. For let G', G'' be the apparent perveance at the voltages e', e'' from the definition we have

$$G'/G = (e'/e_0)^{\alpha - \frac{3}{2}}$$

and

$$G''/G = (e''/e_0)^{\alpha - \frac{3}{2}}.$$

Hence $$G''/G' = (e''/e')^{\alpha - \frac{3}{2}}$$

or $$\alpha - \frac{3}{2} = \frac{\log(G''/G')}{\log(e''/e')}.$$

As a numerical example take the data $e' = 20$ volts, $i' = 2{\cdot}72$ milliamperes, $e'' = 26$ volts, $i'' = 11{\cdot}5$ milliamperes.

Then $$G' = 2{\cdot}72 \times 10^{-3} \div 20^{\frac{3}{2}} = 3{\cdot}05 \times 10^{-5}$$
$$G'' = 11{\cdot}5 \times 10^{-3} \div 26^{\frac{3}{2}} = 8{\cdot}75 \times 10^{-5}.$$

Then $$\log(G''/G') = 0{\cdot}4579$$

and $$\log(e''/e') = 0{\cdot}1142.$$

Hence $$\alpha - \frac{3}{2} = 0{\cdot}4579 \div 0{\cdot}1142$$
$$= 4{\cdot}0$$

and $$\alpha = 5{\cdot}5.$$

The value obtained on drawing a straight line fitting the points as well as possible is 5·1. It is of interest to compute also the value of G the true perveance. Taking $e = 10$ volts and $i = 0{\cdot}6$ milliampere we have

$$G = 0{\cdot}6 \times 10^{-3} \div 10^{\frac{3}{2}}$$
$$= 1{\cdot}9 \times 10^{-4}.$$

Influence of Gas on Saturation Current.

88. In a perfect vacuum the value of the saturation current is equal to the total number of electrons emitted per second from the filament, but in a tube containing gas there may be some current due to positive ions. It is found that in the modern ionic tubes this augmentation is inappreciable unless unusually high voltages be applied. We may conclude from this that in the ordinary way of using a tube all the positive ions are neutralised by slow moving electrons encountered in the dense portion of the space-charge near the cathode, and that the saturation current cannot exceed the emission.

The Inert Gases.

89. The emission from tungsten is almost unaffected by the presence of the gases argon, helium, hydrogen, nitrogen, mercury vapour and certain of the rarer gases. This experimental fact has already been taken for granted in Fig. 250. The main influence

of the gas is to make saturation occur at a lower voltage than in a vacuum, for instance, at 20 volts instead of 60. This is true even at the relatively high pressure of 100 bars in the case of argon. The saturation part of the characteristic in the case of mercury is even flatter than might be obtained in a very good vacuum, and this although the voltage is raised so high that the ionisation by collision gives rise to the blue glow.

Nitrogen.

90. Nitrogen has no effect on the emission from tungsten when saturation is reached at a low voltage without much ionisation, but when the voltage is raised sufficiently nitrogen

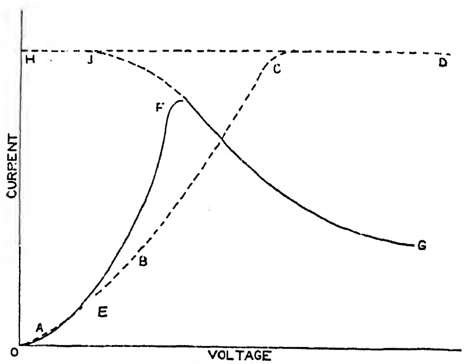


FIG. 253.—Characteristic of diode containing nitrogen.

molecules are driven on to the filament and caused to combine with the metal. The result is that as the voltage is raised the filament gets coated with nitrogen molecules or with a compound and the emission falls off. For example, a tube containing nitrogen at a pressure of 2·5 microns gave a saturation current of 0·75 mA at 70 V; but at 240 V the current was only 0·40 mA. This is indicated in Fig. 253 by the part of the characteristic marked EF. At any temperature of the filament and at any constant voltage the nitrogen compound is continually distilling off as fast as formed, and therefore if the voltage is lowered the filament regains its emitting properties. Quick changes of voltage exhibit a time lag and produce a looped curve. Sometimes there is tungsten on the glass walls of a bulb and then

positive ions of nitrogen driven into the metal never escape again. The consequence is that the nitrogen more or less rapidly disappears.

Oxygen. Water Vapour.

91. Oxygen has a profound influence in reducing the emission from tungsten, even when the pressure of the oxygen is a million times smaller than that at which argon is still inert. Water vapour at very low pressure such as a thousandth of a micron has the same effect as pure oxygen. In both cases it would appear that a layer of oxygen molecules, perhaps combined as tungsten oxide, forms on the tungsten surface, and, since oxygen

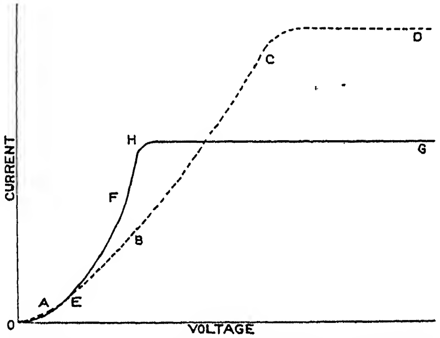


FIG. 254.—Characteristic of diode containing oxygen.

is less electro-positive than tungsten, reduces the emission. A layer one molecule thick is sufficient to modify the vacuum characteristic in the manner indicated in Fig. 254. If the gas admitted for the purpose of experiment is pure dry oxygen it is found that in a time depending on the filament temperature the tungsten oxide distils to the walls of the bulb and then the full emission is permanently regained. But if the oxygen is admitted as water vapour the hydrogen liberated when the oxide forms on the filament can combine with and reduce any tungsten oxide on the walls of the bulb so as to reform water vapour. This vapour then travels to the filament and goes through the same reaction as before. The process is therefore endless and the effect of water vapour is thus a permanent one.

Other Gases.

92. An interesting characteristic given, like the previous ones, by Langmuir, is that of Fig. 255, which depicts the properties of a diode containing argon with a trace of water vapour. The saturation curve FH is produced backward to E and represents the emission from zero voltage to the highest voltage. The ordinate OE is much less than in vacuum and indicates the amount of contamination of the filament by oxygen. As the voltage rises the increasingly vigorous bombardment of the filament by heavy argon molecules gradually clears it of contamination and so improves the emission. The clearing may be either an actual

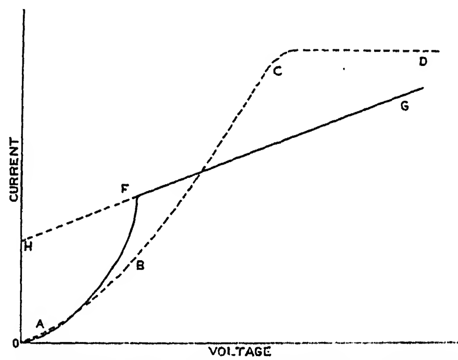


FIG. 255.—Characteristic of diode containing argon and water.

removal of oxygen molecules or a sputtering of tungsten carrying water molecules to the glass.

Carbon monoxide, methane and various hydrocarbons react with incandescent tungsten and leave a deposit of carbon on its surface. Since the emission from carbon is about a third of that of tungsten, the total emission may become much reduced.

93. It is evident from the above paragraphs that the mere heating of a filament in an ionic tube may cause all sorts of changes. Hydrogen is dissociated into atoms, oxygen combines with tungsten, nitrogen and carbon monoxide react with tungsten vapour, hydrocarbons are decomposed and carbon deposited on the filament. At higher temperatures carbon and tungsten oxides distil to the walls. The gas pressure is always changing when a tube is in use, but as a rule there is a progressive reduction of the gas pressure which is roughly proportional to the emf

applied between the electrodes. This phenomenon is called the "clean-up." Under any given circumstances the clean-up proceeds until an equilibrium is reached between this process and the slow release of gas from the electrodes and the glass.

94. The chief source of hydrogen and water vapour in an evacuated bulb is the glass. Glass usually contains a relatively large amount of water, but the bulk of it can be removed from the surface layers by heating the bulb to 380° C. during exhaustion. The more effective way is to use a vacuous oven in which the evacuated bulb can be heated nearly to the softening point without collapsing. If these precautions have been neglected in the manufacture the mere heating of a filament in the bulb provokes copious and long-continued evolution of gas. This gas cannot come from the filament, for the same result is obtained if the filament used has been previously "flashed" to a temperature of 2700° K. in high vacuum, which is known to deprive it of practically all its gas.

The evolution of gas from glass during the use of the tube is minimised by cooling the bulb in an oil bath or by equivalent means, and by strong cooling it is even possible to absorb gas into the glass and so prevent the vacuum deteriorating.

Indirect Influence on the Saturation Current.

95. When electrons or positive ions which have been set into motion in an electric field fall upon an electrode they give to the electrode their electric charge and their kinetic energy. The mechanical aspect of the matter is called "bombardment," as has already been explained in § 78. Bombardment by positive ions is, however, not so simple a process as by electrons. According to § 88, positive ions do not in general reach the cathode un-neutralised, yet it is well known that a filament may be disintegrated rapidly in such a gas as mercury vapour or in argon. Evidently a positive ion may acquire momentum in the field, may pick up an electron and become neutral, and then, in virtue of its inertia, continue in the line of motion of the ion and ultimately reach the cathode and deliver its momentum as a blow. This is sometimes called "bombardment by positive ions," but would perhaps be more correctly styled "bombardment by molecules." This bombardment may influence the saturation curve in several ways.

96. In the first place, bombardment of the cathode heats it, enhances the emission and causes an apparent increase of the saturation current as the emf is raised. This increase of current causes further heating and emission and may produce a large effect cumulatively. In the second place, bombardment may alter the surface of the filament and therefore modify the emission; this has already been discussed as it applies to gas films on a filament, but other effects are possible. The case of a thoriated filament is important. In general, a speck of thoria on the surface of the tungsten, being more electro-positive than the tungsten, attracts the positive ions towards itself and suffers accelerated bombardment, which produces more heating and emission and still greater bombardment, with the result that the speck of thoria is quickly removed by even low pressure gas. This local heating may in gas of relatively high pressure lead, by the cumulative process just described, to the establishment of an arc and the destruction of the filament. In any case, local heating makes saturation difficult to obtain and renders determination of the average temperature impossible. It is noteworthy that the bombardment of pure tungsten, even if so large as to be very destructive, does not in itself affect the emission.

97. Another anomaly in the saturation current sometimes comes from the addition of the current across the vacuum to the current in the filament. The question has been discussed for the pure electron discharge in § 44; when gas is present the currents are larger and therefore the effects more marked. The net current in one part of the filament is increased greatly by the anode current and in the other part is decreased. A non-uniform temperature distribution is thus set up along the filament and the saturation curve modified. As in the cases of preceding paragraphs this may lead cumulatively to overheating of one-half of the filament.

Modifications arising in Practice.

98. In the discussion of the motion of the electrons and positive ions we have supposed that these particles confined their wanderings to the space containing the shortest paths from anode to cathode. But in fact the ions in this space are moving with the usual molecular velocities in random directions, and it is clear that many may escape from the space between the elec-

trodes. It is indeed demonstrable that many of the ions make long circuitous journeys that reach into remote regions of the bulb. The clearest evidence is afforded by the visible glow that occurs in all parts of a tube containing a relatively large quantity of gas when sufficient voltage is applied between the electrodes, or, again, by the production of phosphorescence and heat by positive bombardment at points of the glass wall distant from the electrodes. The absorption of gas into all parts of the wall, already alluded to, which takes place at all voltages above the ionising potential, shows that positive ions are repelled from the anode in all directions.

The Third Body Effect.

99. It has been pointed out in § 61 that when the electrodes in a perfectly evacuated diode are well separated the glass, or, indeed, an insulated conductor, may become negatively charged to about two volts relative to the cathode. This happens even when the anode is at zero potential and the electron current very small, for the hot cathode emits a certain proportion of high-speed electrons able to escape against the field of the space-charge. The negative charge on the glass produces at the cathode an electric field that assists the space-charge to push electrons back into the cathode and thus, in such a tube, the effective perveance may be very small. This has been called the third body effect. We have now to consider how it is modified by the presence of gas; we shall confine our attention in this paragraph to the case in which the third body is the glass of the bulb.

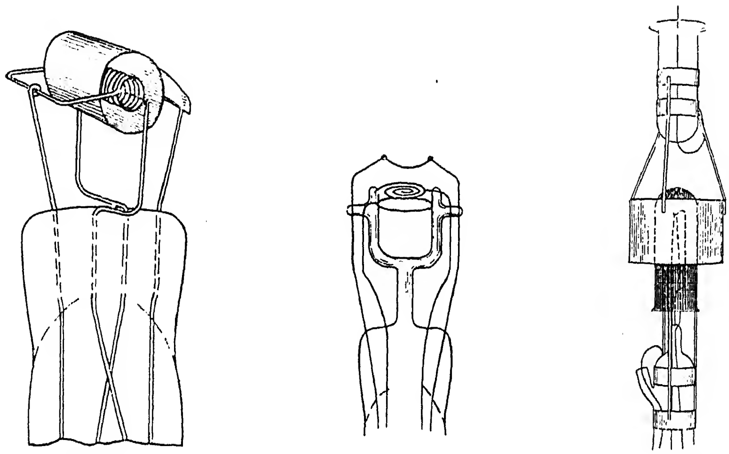
When the anode is negative or a fraction of a volt positive relative to the cathode a certain number of the electrons with high initial velocities reach the glass as before, but the charging process now takes a little longer than in a perfect vacuum because of collisions with molecules; and when the charge is fully established the third body effect is the same in gas as in a vacuum. But when the ionising potential is reached and passed the consequences depend on the pressure of the gas. If the gas is very rare and the voltage only a little greater than the ionising potential the positive ions are few, are all formed very near the anode, and are neutralised by electrons before moving far from the anode; but further raising of the voltage expands the region in which

the ionising velocities are attained and so increases the number of ions formed and their average distance from the anode, until at length the positive ions are attracted to the glass in sufficient numbers to neutralise the negative electricity there and even charge the surface positively. The original negative electric field of the third body is thus reduced or reversed, the space-charge field at the cathode is in some degree cancelled, and the current to the anode rises in value. This is all in addition to the direct effect of the positive ion space-charge upon the electron space-charge.

100. The amount of ionisation increases with the density of the gas in the region where ionising velocities are attained, and therefore, if the gas pressure is higher than in the case just supposed, positive ions will reach the glass in the same numbers at a lower anode voltage than before. If the gas density is supposed greater still the positive ions may reach the glass in such quantity as to charge it positively nearly to the voltage of the anode, and the perveance of the tube may be enormously increased; but now other actions may take place. In some conditions the positively charged glass attracts high-speed electrons, the glass becomes hot, fluoresces and emits secondary electrons as described in § 61. The conditions are not easily controlled, and therefore the characteristics of the tube are found to undergo erratic changes following slight changes of voltage, or even on bringing a conductor near to the outside of the bulb. The course of events depends greatly on the form of the diode. When the distance separating the electrodes is larger than the distance from an electrode to the glass the third body effect is very prominent; but when the anode is, for example, a long narrow cylinder, the cathode an axial filament, and the bulb much larger than the electrodes, the movements of the ions are largely confined to the interior of the cylinder; and in any case the glass, even if it becomes highly charged, can have little influence on the electric field round the filament.

TRIODES.

101. In the discussion of diode tubes it has been necessary to pay some attention to the influence of any third body capable of being electrified and near enough to the electrodes to affect the field of the space-charge. In triodes the properties of such



(French.) (German.) (British.)

FIG. 256.—Types of triode construction.

a third body are intentionally made use of. The third body then consists, as a rule, of a "grid" fixed between cathode and anode and connected to a wire leading outside the tube. The term "grid" includes any conducting partition, network or cage, having holes or meshes that permit a large proportion of the electrons to travel straight from cathode to anode; ideally it should offer no obstruction to the transport of electrons. Several types of grid are seen in Figs. 256, 257. But other forms of third conductor have been proposed. For example, a serviceable triode can be made of a rectangular plate as anode, a parallel filament as cathode, and another plate parallel to the anode and on the other side of the filament, as third body. Instruments have also appeared in which the third body was placed outside the glass of the bulb, as shown in Fig. 258. In all

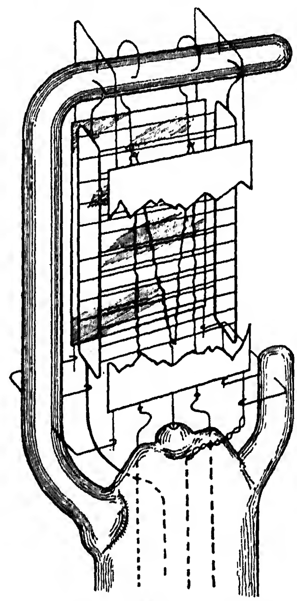


FIG. 257.—Western Electric type K triode. 600 V 0·04 A. Filament 1·5 A.

cases, as used in practice, when the potential of the third body is altered, the field in the space charge near the cathode is correspondingly modified and more electrons escape to the anode, or more are sent back into the cathode, as the case may be; that is to say, the anode current is controlled by means of the potential of the third body. Hence the third conductor is frequently called the control electrode.

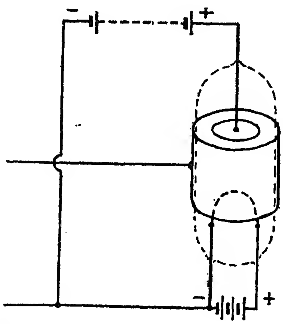


FIG. 258.—Ionic tube with three electrodes. Control electrode outside.

102. The use of the control electrode in the manner indicated was first described in 1907 by Lee de Forest, who introduced a zigzag wire between the filament and plate of the original diode audion. The name "audion" has been carried over to the tube with three electrodes by de Forest and by the Western Electric Company. Later Lieben and Reisz introduced a perforated metal plate between cathode and anode as a control electrode. Then H. J. Round placed grids of gauze in Fleming "valves," and I. Langmuir and his associates introduced the wound wire grid and called the tube a "pliotron." Later, Peri and Biguet produced a tube in which the control electrode is a helix of wire surrounding the

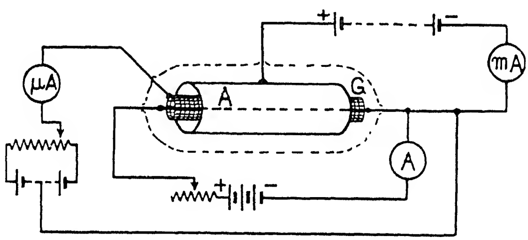


FIG. 259.—Ionic tube with three electrodes. Control electrode inside.

filament. The different names used by various makers are convenient when distinction is to be drawn between them, but since the essential point is the co-existence of three electrodes they will here all be called "triodes." A typical triode and its connections are seen in Fig. 259.

103. It may be remarked that a control electrode may be employed in at least two ways in ionic tubes. In one of these ways the third body when electrified attracts or repels electrons in flight so that they are deflected from the course they would have pursued in the absence of the control; in the other mode of operation the electrified third body accelerates, positively or negatively, the electrons, and so helps them towards the anode or pushes them back into the cathode. In the one case the disturbing action is lateral and in the other is along the line of motion. It is the latter process that is most used in the triodes we have to discuss, at least when the anode voltage is high compared with the simultaneous control voltage.

General Theory of Control Electrode.

104. Before going into further detail in respect of the properties of triodes as elicited by experiment, it is advantageous to investigate theoretically the functions of the electrodes and the space-charge. The results arrived at will assist in the interpretation of experiments, and will suggest methods of measurement and of recording the properties of triodes.

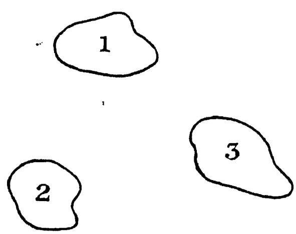


FIG. 260.

Consider three conductors in space, as indicated in Fig. 260, and suppose an electric charge given to each. We shall suppose that the conductor marked 1 is raised to a much higher electric potential than either 2 or 3, and that we wish to inquire how the charge on 2 is influenced by the voltages between 2 and the other bodies. We know from elementary electrostatic theory that q_2, the charge on conductor 2, is given by

$$q_2 = c_{21}v_1 + c_{22}v_2 + c_{23}v_3$$

where v_1, v_2 and v_3 are the absolute electrical potentials of the three conductors, and c_{21}, c_{22} and c_{23} are what are called "capacity coefficients." The properties of these coefficients are set forth fully in A. Russell's "Alternating Currents," Chapter IV. Let us suppose that the potential of 1 relative to conductor 2 is maintained by a connecting wire passing from 2 through a

battery of voltage e_1 to conductor 1, and that a similar wire and battery of voltage e_3 connects 2 to 3. Then

$$v_1 = e_1 + v_2 \text{ and } v_3 = e_3 + v_2.$$

The connecting wires and batteries are assumed to have capacity quite negligible compared with that of any one of the conductors 1, 2 or 3. The above equation becomes

$$q_2 = c_{21}e_1 + c_{23}e_3 + (c_{21} + c_{22} + c_{23})\, v_2.$$

If we take v_2 as the zero of electric potential for our problem we have

$$q_2 = c_{21} \left\{ e_1 + (c_{23}/c_{21})\, e_3 \right\}.$$

We shall write ν for the ratio c_{23}/c_{21} and then the equation becomes

$$q_2 = c_{21}\,(e_1 + \nu e_3).$$

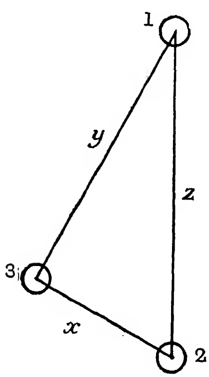


Fig. 261.

This equation shows that when e_1 is increased by, say, one volt, the charge q_2 can be maintained at its old value by reducing e_3 to the extent of one νth part of a volt. In other words, as regards effect on conductor 2, one volt on 3 is worth ν volts on 1. The number ν may or may not be greater than unity—that depends on the electro-geometrical constants c_{12} and c_{23}, which are sometimes called "mutual capacitances."

105. In order to make the matter less vague we may take a simple case and perform the necessary calculations. Let the three conductors be spheres of radii r_1, r_2, r_3, separated by distances x, y, z, much greater than the radius of the largest sphere, as represented in Fig. 261. By a very fundamental principle of electrostatics the absolute electric potential of any sphere, say number 1, is made up by algebraic addition of the scalar potential due to the charge on itself, of that due to the charge on 2 and that due to the charge on 3.

Therefore $$v_1 = \frac{q_1}{r_1} + \frac{q_2}{z} + \frac{q_3}{y}$$

similarly $$v_2 = \frac{q_1}{z} + \frac{q_2}{r_2} + \frac{q_3}{x}$$

$$v_3 = \frac{q_1}{y} + \frac{q_2}{x} + \frac{q_3}{r_3}.$$

On solving these equations for q_2 we find (on neglecting a constant factor not needed in our present application) that

$$q_2 \propto -\left(\frac{1}{zr_3} - \frac{1}{xy}\right) v_1 + \left(\frac{1}{r_1 r_3} - \frac{1}{y^2}\right) v_2 - \left(\frac{1}{xr_1} - \frac{1}{yz}\right) v_3.$$

The coefficients of v_1, v_2, v_3 are capacity coefficients of the system, and by comparing with the general equation for q_2 above we see that in our present problem

$$\nu = \left(\frac{1}{xr_1} - \frac{1}{yz}\right) \div \left(\frac{1}{zr_3} - \frac{1}{xy}\right)$$

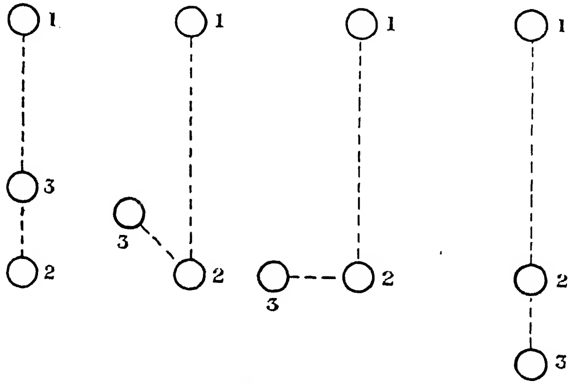


FIG. 262.

To prepare for a numerical example take $r_1 = r_3 = r$, that is, let the spheres be of equal radii, then

$$\nu = \frac{yz - rx}{xy - rz} \doteqdot \frac{yz}{xy} \doteqdot \frac{z}{x}.$$

This result shows that when the conductors are small compared with the distances between them the value of ν is equal to the ratio of the distances of 1 and 3 from 2, and is independent of the distance of 1 from 3. For example, if $z = 3x$, as in all the cases of Fig. 262, the value of ν is 3, and, therefore, one volt applied between 3 and 2 is worth 3 volts applied between 1 and 2, as regards effect upon the charge on 2.

106. It is convenient to remark here for purposes of future reference that the three conductors may be replaced for many electrical problems by a network of ordinary " closed " condensers.

This is indicated in Fig. 263. The leads connected to the condenser plates are supposed to have no capacitance and the surface of zero potential is, in practice, the walls of the room surrounding the three bodies. Then, by the ordinary rule connecting the P.D. of the plates of a condenser with the magnitude of its positive and negative charges, and assuming that 1 is at the highest potential and 2 at the lowest, we have, using the symbols of the diagram,

$$\begin{aligned} q_2 &= C_{22}v_2 - C_{12}(v_1 - v_2) - C_{23}(v_3 - v_2) \\ &= -C_{12}v_1 + (C_{12} + C_{22} + C_{23})v_2 - C_{23}v_3. \end{aligned}$$

Comparing this with the equation involving capacity coefficients we obtain

$$c_{12} = -C_{12},\ c_{23} = -C_{23}$$
$$c_{22} = C_{12} + C_{22} + C_{23}.$$

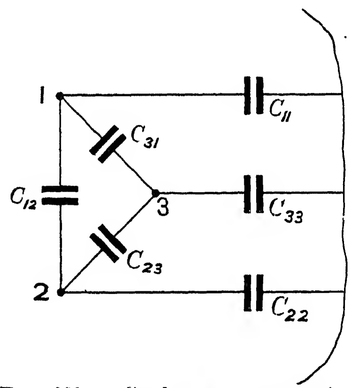


Fig. 263. — Condensers representing capacitances in a triode.

Similar and consistent equations are obtained on writing down the equations for q_1 and q_3 and the system of conductors thus represented as a group of connected condensers. Simplifications can evidently sometimes arise, as, for instance, when two of the bodies are practically enclosed in the other; for if 2 and 3 are surrounded by 1, for example, the condensers C_{22} and C_{33} may be neglected in comparison with C_{12} and C_{31} and left out of the equations.

107. In all the above discussion the conductor 1 is the analogue of the anode of a triode, 2 corresponds to the cathode and 3 to the control electrode. The charge q_2 on the cathode, being equal to the number of lines of electric force ending on the cathode, may be taken as the measure of the average electric field there. All this applies strictly to a triode when cold; but when the cathode is emitting electrons the presence of the space-charge gives vagueness, so to speak, to the boundary of the conductor, and, besides, the dynamics of the electrons ought to be taken into account. Apart from these divergencies, which are found not to be very important, an approximate theory of a small con-

trol electrode in a triode could be completed by superposing the potential distribution due to the three conductors on the distribution due to the space-charge that would be formed in the absence of the control electrode.

Theory of the Grid.

Triode with Plane Electrodes.

108. We shall first suppose the cathode and anode to be parallel planes of infinite extent and the grid to consist of a number of equidistant straight parallel wires with their axes in

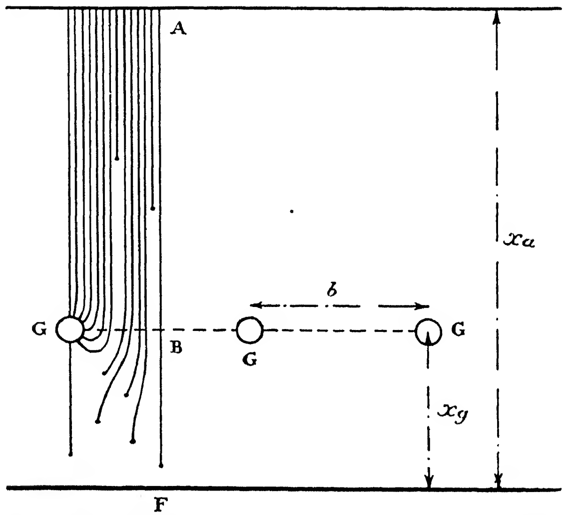


Fig. 264.—Diagrammatic representation of lines of force from anode ending on grid wires and on flying electrons.

a plane parallel to the electrodes. The cathode will be supposed to be emitting the electrons uniformly over its surface. It is possible to attack the problem by a method based upon the calculation of capacity coefficients, making allowance afterwards for the influence of the space-charge upon the electric field between the electrodes, or we could make use of Maxwell's well-known solution for the shielding action of a grid of equal, parallel, evenly spaced wires placed between and parallel to a plane anode and cathode (Maxwell's "Electricity and Magnetism," p. 248, first edition), but it is more instructive to make a direct attempt upon the problem.

We may begin such an attempt by endeavouring to picture the instantaneous state of the electric field between the electrodes. The form of the lines of force is suggested in Fig. 264. Since the triode is usually used with the anode at much higher potential than the grid and the cathode, the Faraday lines are portrayed as starting on the anode A and ending on the grid G and on the electrons of the space-charge. Near the anode the density of the lines is uniform and the lines are perpendicular to the anode. They remain nearly parallel till near the grid wires, and here the radial field due to the charge on the grid combines with the parallel field to give the curved lines. Passing into the region between the grid G and the cathode F the field becomes uniform again, but is clearly much weaker on account of many of the lines having been disposed of. None of the lines reach the cathode because the field, as we have already seen, must be zero at or near the cathode ; that is to say, all the lines that penetrate the shield formed by the grid end upon flying electrons. In the figure there is marked a region B wherein the grid exerts very little influence. We shall make this a kind of stopping place in our calculation of the voltages between different points in the field.

109. Let x_a represent the distance of the anode from the cathode, x_g the distance of the plane of the grid from the cathode, b the distance between consecutive grid wires, r the radius of each grid wire, and e_a, e_b, e_g the potentials of the points A, B, G, relative to the cathode F. The radius r will be supposed small compared with b and the length b small compared with x_a and x_g. We shall now estimate the voltages between the points A and B, B and G, B and F, and shall make these obey the condition that all the Faraday lines leaving A end upon the grid and upon electrons. In order to do this without great labour it is necessary to make some simplifying approximations at the very beginning. First we shall assume that the electric field is practically constant all the way from A to B ; then if σ_a be the number of lines starting from the square centimetre of anode, the field is $(4\pi/\kappa)\sigma_a$, where κ is the inductivity (S.I.C.) of the vacuum, and the P.D. between A and B is

$$e_a - e_b = 4\pi\sigma_a(x_a - x_g)/\kappa$$

Next let the charge on every centimetre of each wire of the grid be Q, then since there are $1/b$ wires per centimetre the

amount of electricity per square centimetre of the grid is Q/b. Now in estimating the P.D. between B and G we may suppose a unit test-charge of electricity moved from B to G under the electric forces due to the whole of the grid wires and calculate the work done. The contribution of each grid wire to the voltage is of the form

$$\frac{2Q}{\kappa} \log \left(\frac{\text{initial distance}}{\text{final distance}}\right)$$

where the distance is to be measured from the centre of the wire concerned. The sum of all such logarithms is of the form

$$\frac{2Q}{\kappa} \log \frac{\beta}{r}$$

where β is a geometrical constant that can be determined by calculation for any number of wires if required. For one wire β is clearly $\frac{1}{2}b$, for two wires, one on each side of B, $\beta = \frac{1}{4}b$, and so on. We obtain, therefore,

$$e_b - e_g = \frac{2Q}{\kappa} \log \frac{\beta}{r}.$$

110. The next step is to evaluate the voltage between B and F. We do this by aid of the electron current equation of § 26, which states that when an electronic current of density i is flowing between two parallel unit plane areas, at one of which the voltage and the electric field are zero, then the voltage at the other is

$$e_b = (i/A)^{\frac{2}{3}} x_g^{\frac{4}{3}}.$$

The last of the equations to be formed is, as already explained,

$$\begin{aligned}\sigma_a &= Q/b + \text{all space charge}\\ &= \alpha Q/b, \text{ say,}\end{aligned}$$

where α is a quantity to be determined by experiment and obviously approaches unity when the space charge, that is to say, the electron current, is very small. Experiment shows that for most types of triode α is nearly equal to unity for a fair range of values of grid voltages and anode current, but the mode of derivation shows that it should tend to increase as the electron current increases.

111. We now have four equations amongst the seven variable quantities and can therefore eliminate three of them, say σ_a, Q and v_b. From the first equation we obtain σ_a, from the second Q; substituting these in the fourth equation we have

$$e_a - e_b = \frac{2\pi\alpha(x_a - x_g)}{b \log \beta/r} (e_b - e_g)$$

or
$$e_a + \nu e_g = (1 + \nu)e_b.$$
where we have put, for brevity,
$$\nu = \frac{2\pi\alpha(x_a - x_g)}{b \log \beta/r}.$$
Now using the third equation we obtain
$$e_a + \nu e_g = (1 + \nu)\,(i/A)^{\frac{2}{3}}x_g^{\frac{4}{3}}$$
or
$$i = A\,\frac{(e_a + \nu e_g)^{\frac{3}{2}}}{(1 + \nu)^{\frac{3}{2}}x_g{}^2}.$$
This formula for the planar triode is not of immediate utility because very few, if any, instruments have been made with the cathode consisting of a heated plane surface. Even in those

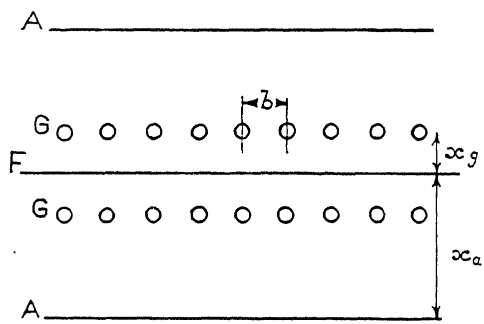


Fig. 265.—Sectional view of triode.

cases where the anode and the grid are planes the cathode is usually a filament, which may in large lamps be wound to and fro between the grid so as to form a number of V's ; and to this the formula is not genuinely applicable. But before leaving it we may compare the equation for ν with that involved in Maxwell's solution of the shielding problem, mentioned in § 108. We then find that for an infinite planar triode
$$\beta = b/2\pi.$$

The Cylindrical Triode.

112. The cathode in a cylindrical triode is a straight filament, or a curled filament with straight axis and of small overall diameter, while the anode will be taken to be a circular cylinder co-axial with the cathode. All the electrodes are supposed to be of infinite length. The grid will be taken to be a collection of equal parallel equidistant co-axial circles, or a co-axial intermediate helix with equidistant spires. In Fig. 265 a sectional view

is given, and A is the anode, F the cathode, G the grid. When, as is usual, the anode is of higher potential than the other electrodes, the Faraday lines starting from the inside surface of the anode form a converging radial electric field in which every line of force ends upon the grid or upon an electron in the space-charge conveying the current. Let the radius of the grid wire be r, the distance between consecutive turns b, the radius of the grid surface x_g, and the radius of the inner surface of the anode x_a. Also let the charge per unit length of the anode be Q_a and that upon unit length of the wire of the grid Q. It follows that the charge per unit length of the grid helix is approximately

$$2\pi x_g Q/b.$$

113. The analysis now follows closely that used in the case of the planar triode. But the potential between A and B is due on the present occasion to a radial field, and the P.D. is therefore a logarithm ; that is to say,

$$e_a - e_b = \frac{2Q_a}{\kappa} \log \frac{x_a}{x_g}.$$

The next equation is, by reasoning the same as before,

$$e_b - e_g = \frac{2Q}{\kappa} \log \frac{\beta}{r},$$

and then, from the properties of the electron current flowing between cylinders, we have

$$e_b = (i/2\pi A)^{\frac{2}{3}} x_g^{\frac{2}{3}}.$$

The condition that the lines of force leaving the anode all end on the grid and on the space-charge may be written

$$\begin{aligned} Q_a &= 2\pi x_g Q/b + \text{space charge} \\ &= 2\pi\alpha x_g Q/b, \text{ say,} \end{aligned}$$

where α is a parameter, to be determined by experiment, usually not much greater than unity.

114. From the first, second and fourth equations we get

$$e_a + \nu e_g = (1 + \nu) e_b,$$

where $$\nu = \frac{2\pi\alpha x_g}{b} \cdot \frac{\log (x_a/x_g)}{\log \beta/r}.$$

On substituting for e_b we obtain finally

$$i = \frac{2\pi A}{x_g} \left(\frac{e_a + \nu e_g}{1 + \nu}\right)^{\frac{3}{2}}$$

where $A = 2{\cdot}33 \times 10^{-6}$ in practical units.

In those cases where the radius of the filament, or of the curled

filament, exceeds about one-tenth of the radius of the grid, it is necessary to apply to the formula just obtained the correction explained in § 32. Ignoring this correction we may write the current through a length l of a triode

$$i_a = \frac{14{\cdot}7 \times 10^{-6} l}{x_g} \left(\frac{e_a + \nu e_g}{1 + \nu} \right)^{\frac{3}{2}}$$

in practical units. It must be recalled that the formulæ are obtained on the supposition that the cylinder, grid and filament are of infinite length, and they cannot be expected to yield very accurate results with the ordinary short triode. It is to be remembered also that the formula for the anode current is obtained under the assumption that the grid current is zero and that the anode is at considerably higher potential than the grid.

115. If the formula for the current per unit length be compared with that deduced in § 32 for the diode, we recognise that the current through a cylindrical triode is the same as that through a cylindrical diode of the same length and of radius equal to that of the surface on which the grid of the triode is wound, provided that the voltage applied to the diode is

$$\frac{e_a + \nu e_g}{1 + \nu},$$

where e_a, e_g and ν relate to the triode. This voltage may be called the "reduced voltage at the grid distance." The quantity $14{\cdot}7 \times 10^{-6}/x_g(1 + \nu)^{\frac{3}{2}}$ may be called the "perveance at zero grid voltage." The quantity $e_a + \nu e_g$, which is especially useful in discussing the relationship of a triode to connected external circuits, will be called the "lumped voltage." Thus the lumped voltage divided by $1 + \nu$ is the reduced voltage.

The current obtained when the grid is connected to the filament so that $e_g = 0$, is proportional to

$$\frac{1}{x_g} \left(\frac{e_a}{1 + \nu} \right)^{\frac{3}{2}}.$$

On the other hand, when the grid is connected to the anode the current is proportional to

$$\frac{1}{x_g}\ e_a^{\frac{3}{2}}.$$

This last result shows that connecting the grid to the anode is equivalent to bringing the anode and its voltage on to the cylindrical surface upon which the grid is wound. The latter expres-

sion divided by the former yields $(1 + \nu)^{\frac{3}{2}}$. This suggests that at values of voltage and current where the three halves power law is obeyed, the value of ν could be obtained by measuring the ratio of the currents through the triode when the grid is connected in turn in the two ways indicated.

Numerical Example.

116. As an example of § 114, a particular reception triode of which measurements have been taken may be quoted here. In this triode $x_a = 0{\cdot}503$ cm, $x_g = 0{\cdot}23$ cm, $b = 0{\cdot}169$ cm, and we may take $\beta = b/2\pi$ for the reason given in § 111. Then, on substituting in the formula, we obtain

$$\nu = 6{\cdot}71\alpha.$$

Now experiment showed that ν is actually equal to 6·9 ; hence we conclude that the quantity α is equal to about 1·03.

Again, the formula for the current yields for an anode voltage of 60 and zero grid voltage the value

$$j = 1{\cdot}31 \text{ mA per unit length}$$

if no account be taken of the radius of the cathode. In fact the cathode is a curly filament of about 0·54 mm radius. The correction factor is seen from the table in § 32 to be about 1·1, which gives

$$j = 1{\cdot}44 \text{ mA}.$$

Actually the measured current was 1·65 mA. This is not a closely accordant result, but is not unsatisfactory when regard is had to the number of circumstances neglected (such as the finite length of the electrodes) in the derivation of the formulæ.

117. A formula for the voltage factor of a different type of triode has been worked out by J. J. Thomson by more rigorous methods. In this the grid consists of a number N per centimetre of straight wires arranged at equal distances apart and parallel to the filament on the surface of an imaginary co-axial cylinder of radius x_g. Thomson's formula is

$$\nu = 2\pi N x_g \cdot \frac{\log (x_a/x_g)}{\log (1/2\pi N r)}.$$

On comparing this with the formula of § 114 we see that in each case the multiplier of the logarithmic fraction is the total straight length of grid wire contained in one centimetre of length of the triode, for by definition $N = 1/b$. On substituting this value of

N and putting $\beta = b/2\pi$, as suggested by the shielding formula of Maxwell, the logarithmic fractions also become identical. Let l_g represent the total length of grid wire per centimetre run of the triode, then the above results show that for either type of grid we may write

$$\frac{\nu}{l_g} = \frac{\log (x_a/x_g)}{\log (b/2\pi r)}.$$

118. This identification enables us reasonably to extend the same formula to include the case of a grid composed of a rectangular network of fine wires and large mesh, for this may be regarded as the result of superposing a grid of circles upon a grid of straight parallel wires. That is, the formula may be applied to a net of fine wires if l_g be taken to represent the total unfolded length of wire woven into a centimetre of the triode. Moreover this identity between the formulæ is sufficient justification for employing the same formula for the type of triode in which the grid is a helix of fine wire.

119. The formula for the voltage factor is worthy of further attention at this point. In the first place it is to be noticed that if all the linear dimensions of a triode be altered in the same ratio the value of ν is unaltered. Secondly, on multiplying the formula for the anode current per unit length by the actual length of the triode, it is seen that the same principle applies to the current at a fixed voltage; that is to say, triodes similar in the geometrical sense have equal perveance as well as equal voltage factor. Next it may be noticed that if the ratio of the anode radius to the grid radius be the same as the ratio of the grid spacing to the circumference of the grid wire the logarithms are equal and $\nu = l_g$. Now this relationship is, as it happens, fairly nearly approached in many commercial reception tubes, and therefore the equation $\nu = l_g$ is often good enough for practical purposes. Further, when this relation is not quite satisfied, we may regard the logarithmic fraction as a correcting factor for the equation. The character of this correction is seen by writing γ temporarily for the correcting factor, so that

$$\gamma = \frac{\log (x_a/x_g)}{\log (b/2\pi r)}$$

or

$$\frac{x_a}{x_g} = \left(\frac{b}{2\pi r}\right)^{\gamma}.$$

Then on putting different constant values in turn for γ we obtain the family of curves sketched in Fig. 266, where values of one ratio are taken as abscissæ and of the other ratio as ordinates. Each curve is marked with the correcting factor to which it

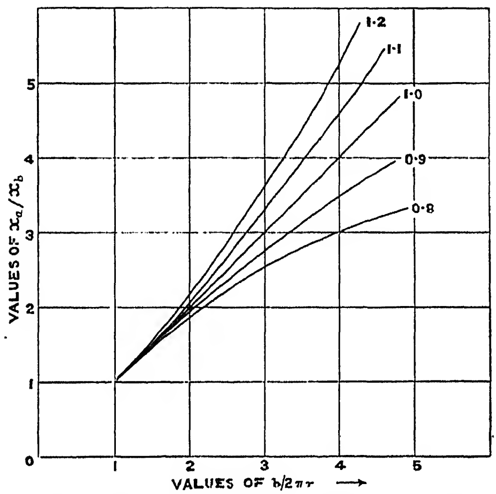


FIG. 266.—Chart for Calculation of Voltage Factor.

corresponds. Alternatively a three-line abac may be constructed to correlate the same variables. Such charts, once drawn, are used by first estimating the two ratios, then finding the point represented by these as co-ordinates, and finally reading off, by interpolation if necessary, the value of the correcting factor which must multiply l_g to give ν.

120. Returning to the form of equation for ν given in § 114 and using Maxwell's value for β we see that the voltage factor is proportional to

$$\frac{x_g \log (x_a/x_g)}{b \log (b/2\pi r)}.$$

It is plain from this that if different values of x_a are taken in turn the numerator alone varies, and since x_a/x_g is usually greater than unity ν increases slowly with increase of x_a. Again, if r be changed the denominator alone varies, and since $b/2\pi r$ is greater than unity the voltage factor increases slowly with increase of r. If, now, x_g alone be increased from a small value the numerator increases to a maximum at the value $x_g = x_a/\epsilon$ where ϵ is the

base of the natural logarithms, and after that decreases. Finally, if b alone be increased from some value greater than $2\pi r$—below which the formula is perhaps invalid—the denominator steadily increases and therefore ν diminishes. To summarise: Within the limits of validity of the formula we find that the voltage factor steadily increases with increase of radius of the anode, with increase of radius of the grid wire and with decrease of the distance between grid wires; it increases up to a maximum as the radius of the surface on which the grid is wound increases up to 0·368 of the anode radius and afterwards decreases.

Limitations of the Electro-geometrical Theory.

121. In applying the above electro-geometrical theory to explain the variations of anode current, the assumption is made that the quantity of electricity released per second from the region of densest space-charge near the cathode is always proportional to the electric field that would be produced near the cathode according to the laws of electrostatics by the charged anode and grid. In the actual operation of the triode the voltage factor is, indeed, made evident only by its property of influencing the anode current. But if conditions arise in which, for any reason, the rate of release of electrons loses its proportionality to the compound electric field, the value of the observed voltage factor will be different from the electro-geometrical value. Extreme cases where this happens are easily conceived. One appears when saturation is attained, for then the anode current neither increases nor decreases when e_g is altered. Again, if e_a is so negative that no electrons can reach the anode, alterations of e_g can have no effect. In short, since in its practical aspect the voltage factor functions by changing the anode current, it is not fully existent when the anode current is zero or non-responsive to changes in the lumped voltage; it is equal to its electro-geometrical value when the anode current is fully responsive, and is smaller than the theoretical value in intermediate conditions.

The quantity called here the "observed value of the voltage factor" is really the ratio of the two gradients discussed in § 135, that is a_g/a_a. If we wished we could logically regard the geometrical voltage factor as a constant and could attribute all the variability of the measured factor to the variability of a_g and a_a.

Filament Potential Drop and Voltage Factor.

122. Next let us consider the effect of the filament drop of potential on the voltage factor ν. Assuming that the anode voltage has a medium value, we may notice first of all that if the grid is made very negative with respect to the negative end of the filament the geometrical theory developed above does not apply fully. When the grid is very negative, and when all the electrons are being returned to the filament, it matters very little whether the grid is made slightly more negative or slightly more positive, as regards effect upon the anode current. Therefore, in measurements of voltage factor by aid of the changes produced in anode current, we expect the working value of ν at negative grid voltages to be much smaller than the geometrical value. As the grid is made less negative the voltage factor increases until at length the grid has reached the potential of the negative end of the filament. At this stage it is obvious that the negative end of the filament will pass more electrons through the grid than the positive end because the latter is confronted by a portion of the grid at lower potential than itself. By the preceding reasoning we therefore expect the voltage factor of the one part of the grid to be greater than that of the other part. When the potential of the grid is raised a little, as in making a measurement of voltage factor, more of the grid becomes positive relative to the filament and correspondingly more of the grid takes the full geometrical value of ν. The aggregate value is what is measured and used, and this may therefore be expected to be greater when the grid is somewhat positive than when it is somewhat negative. Now let the potential of the grid be taken several volts above that of the negative end of the filament. The value of ν should now be the full geometrical value unless saturation is attained at the negative end of the filament. After this happens the aggregate voltage factor will decrease as the grid voltage is increased because more and more of the grid goes out of action.

123. From all this we conclude that the measured value of the voltage factor must depend on the filament voltage for two reasons. Firstly, when the filament voltage is small the potential differences between the grid and the two ends of the filament are less extreme than when the filament voltage is large, and therefore the voltage factor as measured is nearer the ideal in

the former case. Secondly, when the voltage, and consequently the temperature, of the filament is low the saturation of the negative end sets in at a lower potential difference than when the temperature is high, that is to say, the negative end of the triode goes out of action sooner at low filament voltages than at high, and therefore, as the grid voltage rises, the measured value of the voltage factor falls earlier at low filament voltage than it does at high. In Fig. 267 curves appear that show how the

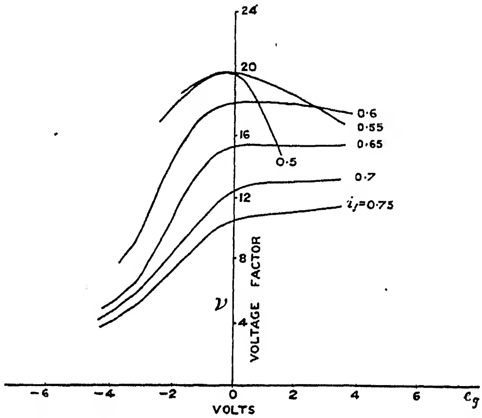


FIG. 267.—Dependence of Voltage Factor on Grid Voltage and Filament Current when measured by Large Changes.

voltage factor of a reception triode varies with the grid voltage and the filament current. The measurements for these curves were made by the method described in § 188. It must be noticed that in all cases something depends on the value at which the anode voltage is fixed; the higher the anode voltage the more negative must the grid be made in order to drive all electrons back to the filament and so reduce the voltage factor to, say, one-half its geometrical value. Again, the higher the anode voltage the less positive need the grid be made to begin the saturation of the negative end of the filament. The curves of Fig. 267 were obtained at an anode voltage of 60 volts. Curves obtained at a higher voltage would fall more quickly on the right hand side of the maximum and less quickly on the left.

Non-uniformity of Temperature and Voltage Factor.

124. The effect of non-uniformity on the voltage factor may

conceivably be considerable. So soon as the ends of the filament become saturated the ends of the grid go more or less out of action and the aggregate value of the voltage factor decreases. The result is to make curves such as those of Fig. 267 descend more rapidly at positive voltages. But it should be noted that adjustments are possible in which an increase of filament temperature acts oppositely to increase of filament emf in respect of alteration of the voltage factor. For when, for instance, only the middle part of a filament remains unsaturated, and only the middle of the grid is in operation for affecting the anode current, an increase of emission all along the filament brings more of the filament into play and therefore increases the value of the working voltage factor.

125. It will be noticed that the variations of voltage factor described in §§ 121 to 124 are mainly associated with rather extreme conditions. In the circumstances in which triodes are usually used these extreme conditions are avoided. In consequence, the voltage factor will be regarded as a constant in much of what follows.

Other Effects of Fall of Potential along Filament.

126. In deriving the equations of the paragraphs preceding § 121 the potential fall along the filament was ignored. We are now in a position to examine this question for the medium circumstances referred to in § 125.

Let us suppose that in making the diagram of Fig. 268 a plane perpendicular to the filament is drawn at any point of it, and that the voltages of these electrodes above that of this point of the filament are represented by ordinates on the corresponding point of FF′. For instance, at the end F of the filament the end of the plate is e_a volts above the filament and the end of the grid is e_g volts above; at the end F′ the plate is $e_a - e_f$ volts above the corresponding point of the filament and the grid is $e_g - e_f$ above, where e_f is the voltage applied to the filament for maintaining its current. In

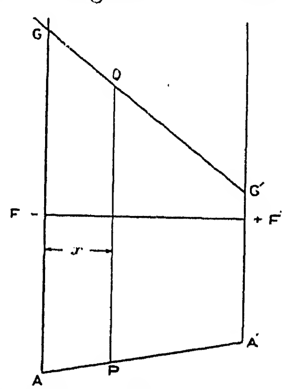


FIG. 268.—Lumped Voltage Diagram.

order that the diagram may at once give the lumped voltage we shall set ν times the grid voltage to scale up the page and the plate voltage to the same scale down the page. This is done at G and G′, and at A and A′. It is evident immediately that the relative voltages in any plane, and the corresponding lumped voltage, are given by joining GG′ and AA′ by straight lines and then drawing an ordinate through the filament abscissa concerned. At a point distant x from the end F the lumped voltage relative to the filament is given by the length PQ. Since the gradient of the anode line is e_f/l, and that of the grid line $\nu e_f/l$, the anode is $e_a - xe_f/l$ volts and the grid $\nu(e_g - xe_f/l)$ "transferred" volts above the filament in the plane distant x from the negative end. The lumped voltage is therefore

$$PQ = e_a + \nu e_g - (1 + \nu)e_f x/l$$

in any plane, and the electron current in this plane is proportional to the three halves power of this expression. The total current at any given adjustment is found by integrating throughout the area AA′G′G. The formulæ for the various cases are easily obtained from § 45 by substituting the reduced voltage, namely, the above expression divided by $1 + \nu$, for the voltage e_a in that paragraph.

127. As the grid voltage rises the line GG′ moves up the page parallel to itself; as the anode voltage rises the line AA′ moves *down* the page parallel to itself. Keeping the anode voltage fixed and varying the grid voltage from extreme negative values to extreme positive values of potential we see that the grid lines take successive positions such as those numbered 1, 2, 3 . . . in Fig. 269. The area swept over by the moving line is at first a rapidly growing triangle, then a rhomboid growing steadily, and then a slowly growing five-sided figure when the line SS′ is intersected. This line represents the saturation level. It is drawn parallel to AA′ because the saturation current is reached at each end at the same relative voltage, that is the voltage AS equals the voltage A′S′.

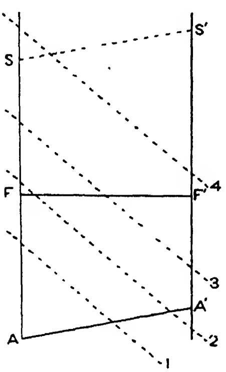


Fig. 269.

128. Inspection of the areas swept over by the moving line shows without elaborate calculation that the effect of the fall

of potential along the filament is practically confined to the bends of the lumped characteristic—the middle portion is not affected—unless the grid is long compared with its diameter, in which case the saturation current may be flowing from one end of the filament before the other end yields any current. This is entered into fairly fully in § 122. The consequence of such a condition is, as has been seen, that the middle portion of the curve connecting the current with the reduced or with the lumped voltage tends to become straight.

Effects of Non-uniformity of Filament Temperature.

129. The effect of the cooling of the ends of a filament, or, indeed, of any lack of uniformity of temperature, have been traced for the case of a diode in § 40. Having now proved that the problems of electron flow in a triode can be reduced to those of a diode, we can accept the conclusions of § 40 with slight change. A principal conclusion was that the non-uniformity of temperature may account for a great part of, or indeed all, the departure from the three halves power law in the middle portion of the curve connecting current and voltage (lumped or reduced). It is interesting to find out in a practical case how far the straightness of a curve is due to the P.D. along the filament, how much to other causes such as non-uniformity of temperature. In the middle curve of Fig. 272 below, the curve is straight while the grid voltage is varied from about -25 to -5 volts, and the anode voltage is 231. The lumped voltage therefore ranges from 58·5 to 196·5 volts, taking $\nu = 6{\cdot}9$. The reduced voltage ranges from 7·4 to 33·6 volts. Now the P.D. along the filament was less than 6 volts. Hence we conclude that in this case the drop of potential along the filament has not been responsible for the straightness of the curve.

Characteristic Curves of the Anode Circuit.

130. Still keeping to the consideration of those modes of use in which the grid current is negligible compared with the anode current, we recognise now that the anode current may be changed, firstly, by varying the anode voltage alone; secondly, by varying the grid voltage alone, or thirdly, by varying both voltages. A curve showing the relation between the anode current and the anode voltage when the grid voltage is constant will be called

the $i_a e_a$ characteristic, and a curve showing the relation between anode current and grid voltage when the anode voltage is constant will be called the $i_a e_g$ characteristic. When both voltages vary together, the relationship between current and voltages can be represented by a characteristic surface, or on a plane sheet of paper by contour lines of constant current, which are sections of the surface by planes parallel to the axes of e_a and e_g. Such curves may be called $e_a e_g$ characteristics. Examples of all these curves will appear on later pages; all of them may be regarded as sections of the characteristic surface which is discussed in § 156.

The Form of the Current Equation.

131. In some respects the most remarkable feature of the mathematical results now reached is the fact that the current through the triode is expressible as a function of $e_a + \nu e_g$. Of course it is obvious without any investigation that the current must be expressible in the form

$$i_a = F(e_a, e_g);$$

but the elucidation of the mode of association of the variables e_a and e_g is of considerable mathematical and physical importance. This point did not wait for theoretical discovery, however; it was found empirically, as already mentioned, by Langmuir (*Proceedings Inst. Radio Eng.*, September, 1915), who showed that large parts of the volt-ampere curves of his pliotrons followed the equation

$$i_a = G(e_a + \nu e_g)^{\frac{3}{2}};$$

and H. J. van der Bijl has also given experimental evidence of the truth of the theorem in the case of triodes of his own design (*Proceedings Inst. Radio Eng.*, April, 1919). This writer prefers to use for the lower values of the currents in his tubes the equation

$$i_a = G(e_a + \nu e_g + e_0)^2.$$

It is possible to fit to a curve any number of equations of the type

$$i_a = G(e_a + \nu e_g + e_0)^n$$

by choosing suitable pairs of values of the parameters e_0 and n, and in such equations the term e_0 will represent the effect of contact difference of potential between the filament and the electrodes, and to a small extent the effect of initial velocities, but Langmuir's equation has the advantage of possessing a physical reason for the appearance of the index 3/2, namely, the

space-charge phenomenon, and in it the perveance G may be regarded as variable. The form $F(e_a + \nu e_g)$ includes both of the above forms, and appears to be applicable over a greater range than either of them; for these cannot hold good so far as the point of inflection in the characteristic curves of i_a plotted with values of e_g as abscissæ. But before discussing these, and before exhibiting any practical applications, it is necessary to point out some of the mathematical consequences of the form of the function.

132. In the first place it is evident that, within limits, any desired current smaller than the saturation current can be produced by a wide variety of values of e_a and e_g; if, for example, $\nu = 20$, then 100 volts on the anode with the grid at zero voltage is equivalent to 120 volts on the anode with -1 volt on the grid, and to 80 volts on the anode with $+1$ on the grid. In fact, we may say that a principal function of a triode is to multiply the grid voltage by ν, to transfer it to the anode circuit, and to superpose it on the anode voltage. We shall give this total voltage, the "lumped voltage," the symbol e_l. The current equation is then

$$i_a = F(e_l).$$

The lumped voltage may now be taken as a new independent variable, and the current curves plotted with values of e_l as abscissæ.

133. Let us take the family of curves in Fig. 270 and transform

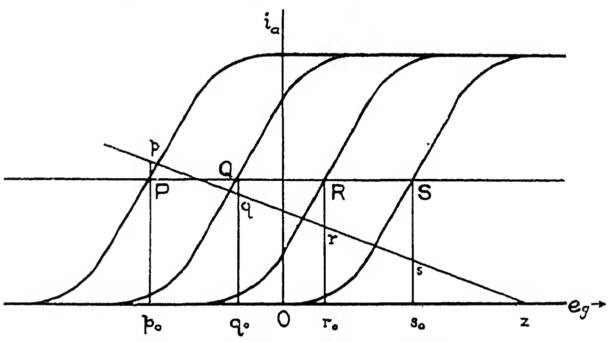


FIG. 270.—Determination of Voltage Factor.

them to the new co-ordinate. It is first necessary to find ν from the curves. An easy way is to draw first any line PQRS to mark points of equal currents on the different curves of Fig. 270;

then set up ordinates p_0p, q_0q, etc., to represent to scale the anode voltage of the respective curves, and draw the best straight line through the points p, q, etc., to cut the axis in z. Since this line relates to constant current, the value of $e_a + \nu e_g$ must be constant along it, and, therefore, a fall in e_a is made up by ν times the increase of e_g. Thus the gradient of the line, reckoned on the voltage scales, is the value of ν. For example, if P is on the 250 volts characteristic and if the length p_0z represents 20 volts on the grid, the value of ν is 250 volts $\div$ 20 volts $=$ 12·5.

134. The next step consists in calculating the value of $e_a + \nu e_g$, that is, e_l, for a number of points on the curves of the family, either arithmetically or graphically. The latter method is shown

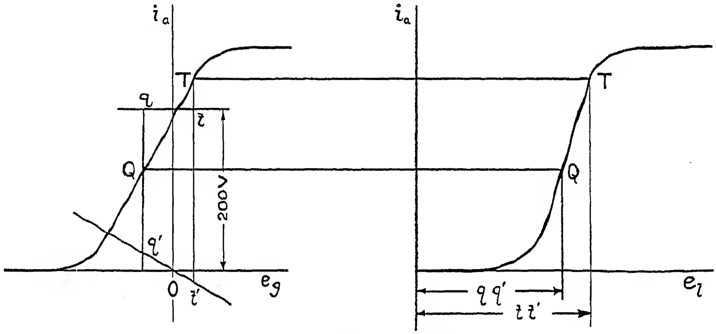


FIG. 271.—Derivation of Lumped Characteristic.

in Fig. 271, where one of the curves is taken aside for the sake of clearness. First, the sloping line through O is drawn parallel to the line through z in Fig. 270. The ordinates of this line are ν times the corresponding grid voltage on a certain scale. Then a level line is drawn to represent to the same scale the anode voltage of this curve (marked 200 V in the figure). If now ordinates are drawn through any points such as Q, T on the given curve, the intercepts between the voltage lines, such as qq′ and tt′, are the algebraic sum of e_a and νe_g, and therefore give values of e_l. The curve sought is now obtained by co-ordinating values of current at Q, T, etc., with the new abscissæ qq′, tt′, etc.; this is carried out on the right-hand side of Fig. 271.

When this process is applied to all the curves of the family in Fig. 272, the result is the nearly perfect single curve of Fig. 273. In this one curve is contained all the information given by the

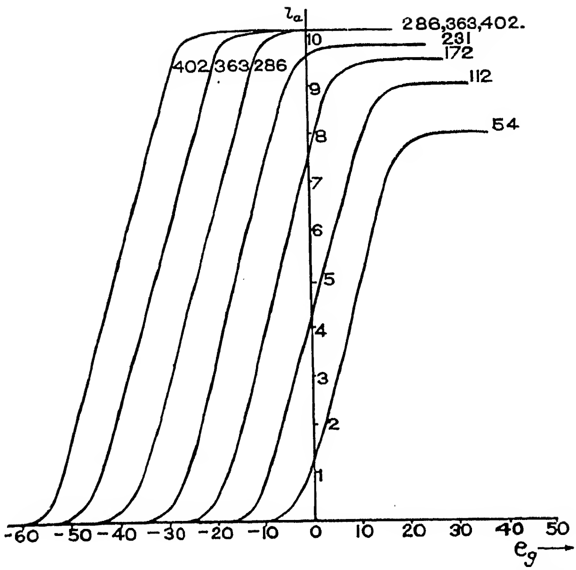


FIG. 272.—Characteristic Curves. Filament Current, 0·75 A.

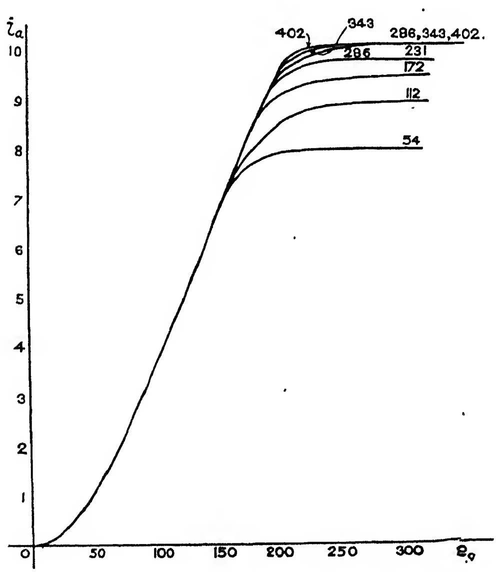


FIG. 273.—Lumped Characteristic of Fig. 272.

original family of curves. It may be called the lumped characteristic curve of the triode for the particular filament current in use. It is found that the nearly straight portion of the $i_a e_g$ curves of most forms of triode coincide perfectly in the lumped characteristic, but the knees fit together less closely. Also, below a certain anode voltage the saturation currents take lower and lower values, with the result that we obtain the forked curve in Fig. 273. The family of curves in Fig. 272 were obtained in the writer's laboratory.

The Gradients of the Characteristics.

135. A further mathematical consequence of the form of our equation is that a simple relation exists between the partial differential coefficients of i_a with respect to e_a and e_g regarded as independent variables. From

$$i_a = F(e_l)$$

we derive
$$\frac{\partial i_a}{\partial e_a} = \frac{di_a}{de_l} \cdot \frac{\partial e_l}{\partial e_a} = \frac{di_a}{de_l}$$

since
$$e_l = e_a + \nu e_g.$$

Similarly
$$\frac{\partial i_a}{\partial e_g} = \frac{di_a}{de_l} \frac{\partial e_l}{\partial e_g} = \frac{di_a}{de_l}\nu.$$

Hence
$$\frac{\partial i_a}{\partial e_g} = \nu \frac{\partial i_a}{\partial e_a}.$$

We shall for brevity write a_a for the partial differential coefficient with respect to e_a, and a_g for that with respect to e_g. In this notation the letter a is meant to suggest "increase of anode current," and the suffix a to suggest "per unit increase of anode voltage," and a similar interpretation is given to a_g. The differential coefficient a_a is, plainly, the gradient at any point of the lumped characteristic curve, and a_g is the gradient at any point of any of the $i_a e_g$ characteristic curves. We see that the latter gradient is ν times the former whenever the current is a function of $e_a + \nu e_g$; but this conclusion is not necessarily true when the current does not possess this property.

Determination of Voltage Factor from Characteristics.

136. It is worthy of note that when the $i_a e_g$ characteristic obeys Langmuir's equation it is possible to deduce the value of ν from a single curve. For, when

$$i_a = G(e_a + \nu e_g)^{\frac{3}{2}},$$

we have

$$a_g = \nu \frac{3}{2} G(e_a + \nu e_g)^{\frac{1}{2}},$$

whence

$$\nu = \frac{2}{3}(e_a + \nu e_g)a_g/i_a.$$

This expression can be evaluated from a given $i_a e_g$ curve by drawing the tangent line at $e_g = 0$ and measuring the subtangent, say OT. Then

$$\nu = \frac{2e_a}{3\mathrm{OT}}.$$

Alternatively the gradient a_g may be determined at any value of e_g, and ν determined from the equation

$$\nu = \frac{2e_a a_g}{3i_a - 2e_g a_g}.$$

The results obtained in this way from reception tubes are usually too small.

137. When a family of $i_a e_g$ characteristics is available, an approximate value of ν may quickly be found by taking any point on one of the family and drawing from it a line parallel to the axis of e_g so as to cut another curve of the family. The length of this line, reckoned in volts according to the e_g scale, gives, when divided into the difference of the anode volts of the two curves, the value of the voltage factor. The reason for this is that the line drawn parallel to the e_g axis represents the condition of constant current, and therefore implies that the change in e_g must be compensated by an opposite change in e_a of magnitude νe_g.

A more accurate way of determining an average value of ν from a family of curves is described in § 133 above.

The Lumped Characteristic and the others.

138. In order to draw the lumped characteristic it is only necessary to apply in turn different known grid and anode voltages and to measure anode current by means of a low-resistance instrument. The lumped voltage is then calculated by multiplying the grid voltage by the voltage factor ν and adding the result to the anode voltage. Methods of measuring ν directly are given below; but, for arithmetical reasons, it is obviously advantageous to use a constant grid voltage, and it is best of all,

from the arithmetical point of view, to use a zero grid voltage. That is to say, it is usually best to connect the grid direct to the negative terminal of the filament, to apply various anode voltages in turn, and then to measure the resulting anode currents. The currents are plotted as ordinates and the lumped voltage as abscissæ, to obtain the lumped characteristic.

139. From the curves obtained we may derive $i_a e_g$ characteristics or $i_a e_a$ characteristics of the kind described in § 130. In order to illustrate the methods of derivation, let us take the curve of Fig. 274 to represent the lumped characteristic of a tube of which it is known that $\nu = 10$. First let us derive an $i_a e_g$ characteristic, say the one for which the anode voltage has a constant value $e_a = 100$ V. To do this, draw a new axis of ordinates at the abscissa $e_l = 100$, and take the foot of this ordinate as the new zero of abscissæ. This operation merely subtracts $e_a = 100$ from every value of e_l, and therefore leaves values of νe_g as the new abscissæ. This is shown by the graduated line drawn below the former base in the figure. In order to obtain an axis graduated in values of grid voltage, we merely divide the numbers of the new base by ν. The result is shown on the base marked e_g. In the figure the graduated axes are shown as separate lines for the sake of distinctness—actually they would all be marked on one line.

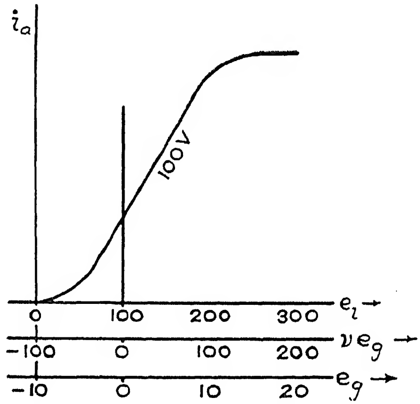


FIG. 274.—Derivation of $i_a e_g$ Characteristic.

The repetition of this process for different anode voltages leads to a series of equal similar curves cutting their axes of current at different heights, and these may, if desired, be all reproduced on one base. We thus obtain the family of curves in Fig. 275. This will be found to agree very well with the results of separate experiments for each anode voltage in all parts of the diagram where the grid current is negligible. The rise of the grid current, which always happens when the grid voltage becomes sufficiently

positive, leads to irregularities such as are shown in Fig. 275 by the dotted line, as has already been remarked. For these regions the lumped characteristic must be supplemented by special measurements.

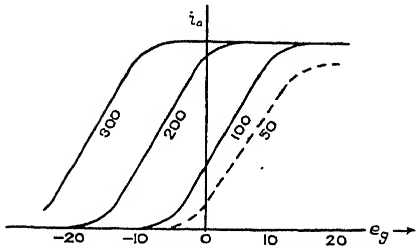


FIG. 275.—Family of $i_a e_g$ Characteristics.

140. Let us now suppose that the $i_a e_a$ characteristic is required for a constant grid voltage of -10 volts. We begin by multiplying this constant voltage by ν, and so obtain -100, which must be subtracted algebraically from every value of e_l in Fig. 274. To do this graphically, move the axis of current to the abscissa -100, as indicated in Fig. 276, and graduate a new base line to the old scale. The original curve is, relative to the new axis, the $i_a e_a$ characteristic required. A family of $i_a e_a$ curves can be obtained from the lumped characteristics by repeating the construction and carrying all the curves to one base line, as shown in Fig. 277. As in the preceding case, we must expect irregularities to enter when positive grid voltages are reached; this is indicated in the figure by the dotted curves, which represent what would be obtained by experiments carried out with the grid voltage at the constant value marked on each curve.

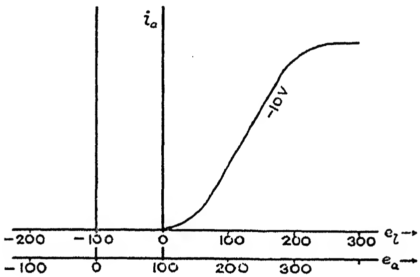


FIG. 276.—Derivation of $i_a e_a$ Characteristic.

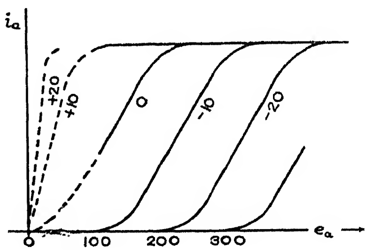


FIG. 277.—Family of $i_a e_a$ Characteristics.

ALGEBRA OF TRIODES.

141. We proceed to apply to the triode considerations similar to those applied to the diode in § 64, keeping still to the simplifying assumption that the

current flowing from the grid to the filament inside the tube is negligible. The electrostatic theory of the control electrode has shown that when it is at a different potential from the filament it superposes an electric field upon that existing in the space-charge, and therefore modifies the back emf due to this charge. When the control electrode is made positive relative to the filament the back emf is reduced by an amount proportional to the voltage of the electrode, which is equivalent to the introduction of a new voltage into the anode circuit and assisting the anode battery. If the potential difference between the control electrode and the filament is e_g, the emf effective in the anode circuit for driving the electron current is now

$$e_a + \nu e_g - e_s,$$

instead of the voltage

$$e_a - e_s$$

occurring in the diode. On giving to the lumped voltage the symbol e_l the theory of the triode is reduced to that of the diode. We have, in fact, along the nearly straight part of the characteristic curve

$$i_a = a_a(e_l - e_s).$$

When it is necessary to express the control voltage explicitly, this may be written

$$\begin{aligned} i_a &= a_a(e_a + \nu e_g - e_s) \\ &= a_a e_a + a_g e_g + a_0, \end{aligned}$$

which is the equation first suggested by M. Latour in *The Electrician*, December 15th, 1916. In this equation $a_g = \nu a_a$ and $a_0 = - a_a e_s$, and all are constants along the straight part of the characteristic.

142. From what has been said regarding the portion of the diode characteristic that satisfies the three halves power law it is obvious that the equation connecting the anode current and the lumped voltage below the heel of the triode characteristic may also be written

$$\begin{aligned} i_a &= a_a(e_l - e_s) \\ &= a_a e_a + a_g e_g + a_0. \end{aligned}$$

In problems where the second differential coefficient of i_a with respect to e_a or e_g is unimportant, the parameters a_a, a_g and a_0 may be treated as constants, and, plainly,

$$\frac{\partial i_a}{\partial e_a} = a_a, \text{ and } \frac{\partial i_a}{\partial e_g} = a_g.$$

In other cases the equation for the curved portion below the heel must be written (neglecting filament drop)

$$\begin{aligned} i_a &= Ge_l^{\frac{3}{2}} \\ &= G(e_a + \nu e_g)^{\frac{3}{2}}, \end{aligned}$$

which is Langmuir's equation for this section of the triode characteristic. For example, if a triode is being used on this lower portion as an amplifier, and not as a detector, we may use the linear equation and regard the parameters as constants; but if the rectifying properties of the valve are under consideration, the Langmuir equation must be used. In this latter case a_a and a_g are still defined by the differential coefficients above, but they are now functions of both e_a and e_g.

It must be noticed that in all cases the parameters are functions of the filament current.

Numerical Example.

143. In a certain tube, when the anode voltage was 150 and the grid voltage 6·5, the anode current was 5 mA. The value of a_a was known to be 4×10^{-5} and $\nu = 10$. Find the equation of the straight part of the lumped characteristic.

In the first place, the lumped voltage of a given point is $150 + 6{\cdot}5 \times 10 = 215$; hence, by substitution in the first form of the equation, we have

$$5 \times 10^{-3} = 4 \times 10^{-5}(215 - e_s),$$

whence

$$e_s = 90 \text{ volts}.$$

The equation of the straight part is therefore

$$i_a = 4 \times 10^{-5}(e_l - 90),$$

where

$$e_l = e_a + 10e_g.$$

In this tube the current co-ordinate of the steepest place was 5·6 mA. The voltage co-ordinate is obtained by use of the $i_a e_l$ equation above; which leads to

$$\begin{aligned} e_l &= 90 + \frac{5{\cdot}6 \times 100}{4} \\ &= 230 \text{ volts}. \end{aligned}$$

Hence, if we make $e_a = 200$ volts e_g will be 3 volts, and if $e_a = 230$ volts, the steepest point of the characteristic will be at zero grid voltage.

Voltage Amplification.

144. An application of the above theory may be made to the problem given in Fig. 278. Here a resistance R is in series with the tube, of which the differential resistance is, say, r_a. The emf in the anode circuit is that of the battery E, the space-charge back emf e_s, and the transferred grid voltage νe_g. The total resistance in circuit is $R + r_a$.

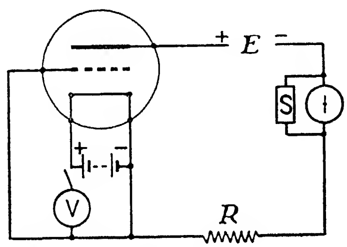


Fig. 278.

Thus $$i_a = \frac{E + \nu e_g - e_s}{R + r_a}.$$

Here e_s and r_a are constants of the tube, supposed known. The current is therefore now determined.

145. In many applications of triode tubes we have to deal with changes of anode current produced by changes of the voltages applied in both the anode and the grid circuits. In this case we have, by differentiation of the last equation,

$$di_a = \frac{dE + \nu de_g}{R + r_a}.$$

For such purposes it is not necessary to know the value of the back emf of the tube. It must be noticed carefully that this theory applies with accuracy only along the straight part of the lumped characteristic of the tube.

Suppose the P.D. at the terminals of R is to be utilised as in an amplifier. Then E is constant and the P.D. is Rdi_a; the voltage amplification is the ratio of this to de_g, namely,

$$\frac{R\nu}{R + r_a}.$$

Geometrical Aspects of a_a and a_g.

146. These quantities, which have been defined in § 135, may be deduced from the equation for the anode current. We have

$$a_a = \frac{\partial i_a}{\partial e_a} \propto \frac{(e_a + \nu e_g)^{\frac{1}{2}}}{x_g\,(1 + \nu)^{\frac{3}{2}}}$$

and $$a_g = \frac{\partial i_a}{\partial e_g} \propto \frac{\nu\,(e_a + \nu e_g)^{\frac{1}{2}}}{x_g\,(1 + \nu)^{\frac{3}{2}}}$$

If, as is usual, ν is much greater than 1 and is (see § 119) approximately equal to $2\pi x_g/b$, we have

$$a_a \propto \frac{b^{\frac{3}{2}}}{x_g^{\frac{5}{2}}}(e_a + \nu e_g)^{\frac{1}{2}}$$

$$a_g \propto \frac{b^{\frac{1}{2}}}{x_g^{\frac{3}{2}}}(e_a + \nu e_g)^{\frac{1}{2}}.$$

Thus the gradients of both the $i_a e_a$ and the $i_a e_g$ characteristics increase when the pitch of the grid increases and its diameter decreases.

Variations of a_a and a_g.

147. In § 146 the values of the gradients of the $i_a e_a$ and $i_a e_g$ characteristics are given. From these results it is seen that both a_a and a_g are proportional to $\sqrt{(e_a + \nu e_g)}$. If therefore they be plotted with values of e_g as abscissæ, taking different fixed values of e_a in turn, parabolic loci are obtained. But evidently in doing this we have neglected the facts that the characteristic becomes nearly straight and then bends over gradually into the saturation region. In the experimental curves of Fig. 279, which are taken from S. Ballantine's paper in the *Proceedings Institute of Radio Engineers*, April, 1919, the left-hand rising portion of each curve is the parabolic locus, the maximum corresponds to the straight part of the characteristic and the downward portion on the right corresponds to the saturation region. All these curves may, if desired, be reduced to a "lumped" curve; and this single curve might be deduced graphically from the lumped characteristics, as indicated in Figs. 272 and 273.

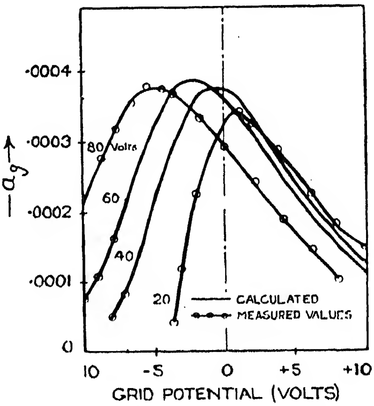


FIG. 279.—Variation of di_a/de_g at various Anode Voltages.

The value of each of these gradients is affected greatly by the conditions of the filament as regards temperature and potential drop; but since in the ordinary use of the tube the grid is

at a voltage much nearer to that of the filament battery than the anode is, the gradient a_g is the more affected by filament conditions. Much of what has been said in § 123 upon the variation of the voltage factor from these causes may be applied at once to the discussion of the variations of a_g and need not be repeated here.

The Grid Current.

148. It has already been pointed out that in a highly evacuated tube an insulated conductor gradually becomes charged negatively by the receipt of electrons, and finally reaches an equilibrium potential whose value depends on the contact potential difference natural to the metals concerned and on the potentials of neighbouring bodies. An insulated grid in a triode is such a conductor, and may acquire a voltage of perhaps a volt below that of the negative end of the filament. If the grid be connected to the filament through a galvanometer, a positive current will in general be observed to flow from the filament to the grid through the galvanometer and will reach its steady value very soon after the making of the connection. Its value is equal to the rate at which electrons can be continuously collected by the grid at the potential it then possesses. This rate depends upon the magnitude of the anode voltage, upon the configuration of the electrodes, upon the temperature of the filament, and also upon the area which the grid offers as a target to electrons supposed travelling in straight lines from the filament to the anode. For example, in a cylindrical triode the area of the target is $2\pi x_g(2r/b)$ per unit length of filament. Reduction of the radius of the wire and increase of the spacing each tend to reduce the grid current. It is of interest to refer to § 120, and to notice that if it be required to reduce the grid current while retaining a certain voltage factor and a certain anode current per unit length, the value of x_g must be retained and only b and r changed; and these changes must be such as to keep $b \log (b/2\pi r)$ constant.

Grid Characteristic, Anode Voltage Constant.

149. Taking a given tube at a fixed filament temperature it is found that the grid current i_g varies with the anode voltage e_a and with e_g. No general theory has yet been given, but, guided by the results of experiments and neglecting the effects of potential drop along the filament, we may exhibit the leading

principles governing the magnitude of the grid current. First assume that the anode voltage is made negative, say about 2 volts. No electrons can alight on the anode and therefore all of them are taken by the grid or returned to the filament. If the grid is insulated it becomes just so negative that no more can alight upon it; but let us suppose that it is connected to the filament through a potential divider so that its voltage may be varied at will. Then when the applied grid voltage is negative the number of electrons collected by the grid will depend upon the distribution of the initial velocities of the emitted electrons

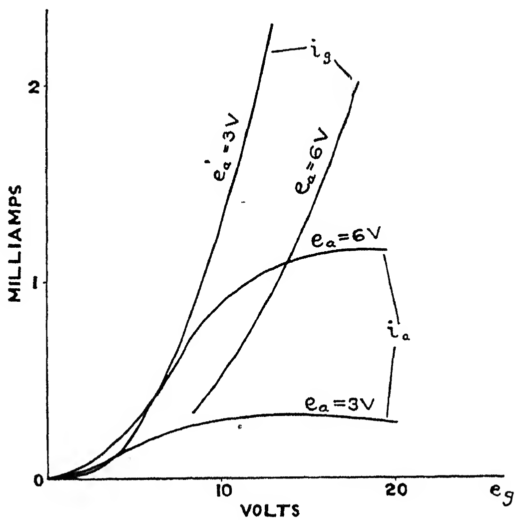


FIG. 280.

as explained in § 36 in discussing the analogous case of the diode; from which we conclude that as the grid voltage becomes less negative the grid current follows an exponential law of the form

$$\log i_g \propto e_g + \text{constant}.$$

When the grid voltage is taken to positive values and gradually increased, the grid collects more and more electrons, and, since we are neglecting the effects of fall of potential along the filament, the grid current will increase according to the three-halves power law. As the grid voltage is increased the grid current increases as the current does in a diode until saturation is attained.

150. Next let the anode potential be fixed at about two volts positive, so that a considerable anode current flows. Begin again

with the grid voltage negative. The most noticeable effect is that the anode current becomes very small, because the negative grid returns many electrons to the filament; but the grid current will increase according to the exponential law as the grid is made less negative. When the grid is at zero potential the current it collects is still small, but on increasing the potential both grid and anode current increase rapidly. Experiment shows that after a certain stage the grid current increases faster than the anode current, the former being finally of greater absolute magnitude. The general shape of the $i_a e_g$ and the $i_g e_g$ characteristics is shown in Fig. 280. Here, however, the filament voltage is considerably distorting the curves from the ideal shape.

151. Usually in the practical applications of triodes the anode voltage is many times greater than the grid voltage, and the

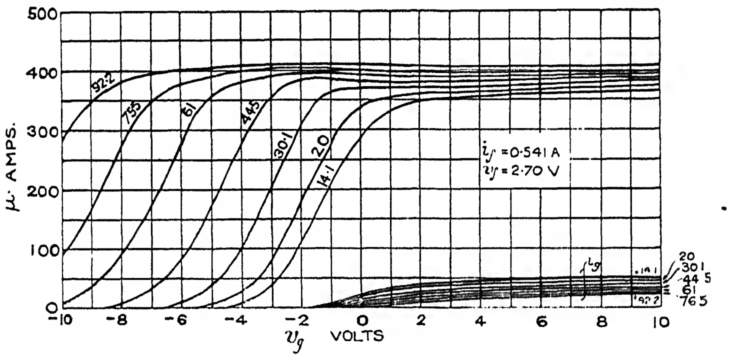


Fig. 281.

electric field near the grid wires is determined principally by the anode. If we suppose the anode voltage about 50 and begin with the grid negative the grid current is again very small. Let the grid voltage now be given a small positive value. The grid current will increase with increase of this voltage but not so rapidly as when the anode voltage was lower, for the anode voltage dominates the motion of the electrons even very near the grid wires so long as it is much greater than the grid voltage. The curves of Fig. 281 show the course of events in some small reception tubes examined by G. Vallauri. In order to reduce these curves to read with the negative end of the filament as the zero of potential it is necessary to move the origin or to add 2·70 to the grid voltage reading.

152. Now let the grid voltage increase until the combined field of the anode and grid draws the saturation current. The ratio of the two currents will depend upon the magnitudes of the voltages, and if the grid voltage be still increased while the anode voltage remains constant we shall find that the saturation current will be shared between grid and anode with increasing bias in favour of the grid. It is thus possible for the grid current to become greater than the anode current in certain types of triode as indicated in Fig. 282, the sum of the two always being equal to the emission from the filament in a perfectly evacuated tube. A pattern of tube that illustrates this phenomenon has been described by J. Erskine-Murray in U.K. Specification 133,413/1918.

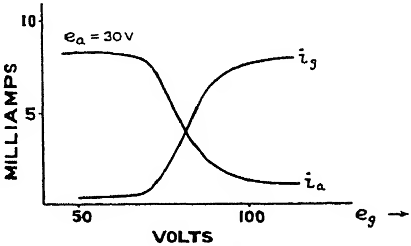


FIG. 282.

Grid Characteristics, Grid Voltage Constant.

153. In contrast with the preceding, it is interesting to trace the result of fixing the voltage of the grid instead of the anode and then varying the anode voltage. As an example, let us suppose the grid voltage at 2 volts, and that we begin with the anode voltage negative and gradually increase it to positive values. When the anode is very negative all the electrons are pushed back into the filament and none can reach the grid, but as the anode becomes less negative the grid current increases, firstly on account of the initial velocities and secondly because the combined field becomes positive and partially neutralises the space charge. For instance, if the voltage factor be 9 and the anode voltage 5, the combined field at the filament is, according to § 115, proportional to $(-5 + 9 \times 2)$, or 13, which is sufficient to cause the liberation of a large number of electrons from the neighbourhood of the cathode. The anode current will remain practically zero until the anode attains the potential of the negative end of the filament, and therefore the grid current will continue to increase until that stage at least. As the anode voltage is further increased into positive values the anode

current will increase by deflecting electrons which might have gone to the grid, hence the grid current falls while the anode current rises; and at length the latter may become of order of magnitude a hundred-fold the grid current. The results of measurements of a small reception tube are shown in Fig. 283. These curves, which were obtained with a filament voltage of about 5 volts, depart perceptibly from the ideal; for on the simple theory outlined above, the sum of the grid and anode currents should be proportional to the three-halves power of

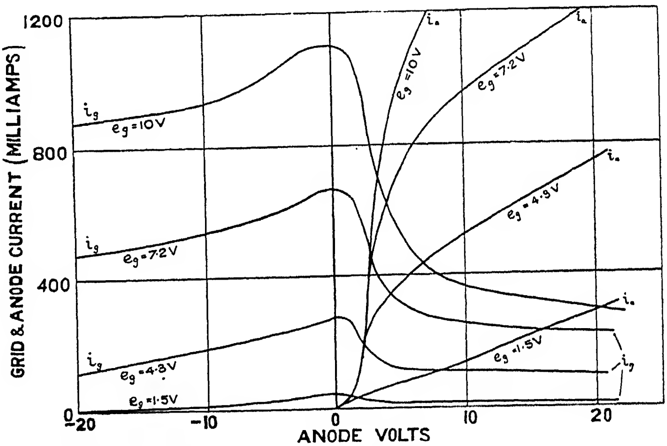


Fig. 283.—Grid and Anode Currents at Low Anode Voltages.

the lumped voltage, and this is not the case. The departure is mainly due to the fall of potential along the filament.

Sharing of Current between Positive Electrodes.

154. In Figs. 282 and 283 the simultaneous rise of grid current and fall of anode current are very striking. This phenomenon is due to the deflection type of action referred to in § 103. In a cylindrical triode some of the electrons starting from the filament in radial directions will pass near enough to a grid wire to be drawn on to it in spite of the attraction of the anode, and of the whole electron stream the proportion so deflected will obviously be larger the greater the ratio of the grid potential to the anode potential. To develop a precise theory of this deflection type of action would take too much space, but it is found that a simplified solution of the equations of motion of electrons passing near

and falling upon the grid wires is expressible for medium and small values of the ratio e_a/e_g by the equations

$$i_g = G\,(e_a + \nu e_g)^{\frac{3}{2}} \exp\,(-\,\alpha\,\sqrt{e_a/e_g})$$

and $$i_a + i_g = G\,(e_a + \nu e_g)^{\frac{3}{2}},$$

where the symbol α is a function of the inverse of the radius of the grid wire. From these equations we deduce that

$$\alpha.\,\sqrt{(e_a/e_g)} = \log\,(1 + i_a/i_g),$$

which may be used for the determination of α by measurements upon a tube. For instance, the family of curves in Fig. 283 yields an average value $\alpha = 1{\cdot}40$ so long as e_a is not greater than $2e_g$,

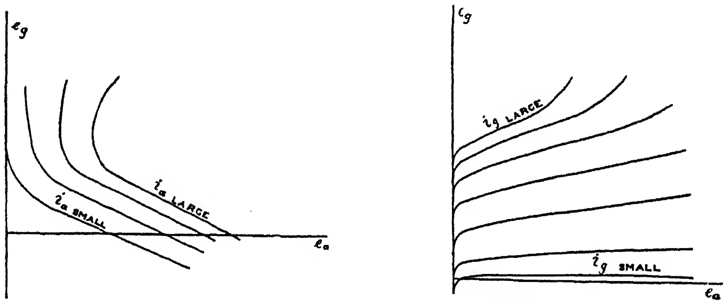


FIG. 284.—Grid Current Contours of Hard Tube.

being rather smaller for greater relative values of e_a and not quite constant.

155. The values of $\partial i_g/\partial e_a$ and $\partial i_g/\partial e_g$, that is to say, g_a and g_g, may be calculated from the above equation, but the results are somewhat too complicated to be useful. The equations are most useful for indicating the shape of the various anode and grid characteristics at low values of e_a and e_g. They agree fairly well with the experimental characteristics already given and lead to the e_ae_g characteristics shown in Fig. 284. These curves are not drawn to scale, but merely indicate the general shape of the contour lines at constant filament current. No attempt has been made to allow for the effect of filament conditions as regards uniformity of temperature or potential, either in the curves or in the equations.

The Characteristic Surface.

156. The total current leaving the region of the cathode has been shown to be a function of the lumped voltage. When the

conditions are ideal the Langmuir equation is followed, but in other cases the equation is less simple and may, indeed, be very complicated; let it be

$$i \equiv i_a + i_g = f(e_a + \nu e_g).$$

Let e_a, e_g and i be measured along a set of mutually perpendicular axes and the surface represented by the equation be drawn. A portion of the surface is seen in Fig. 285 where the plane part corresponds to the saturation current. The surface meets the

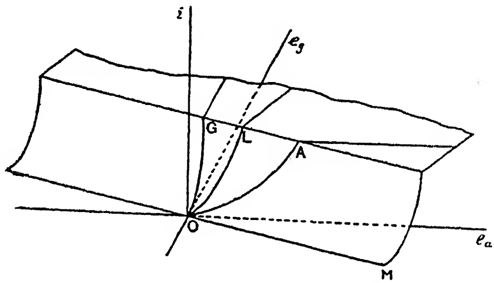


FIG. 285.—Characteristic Surfaces for Total Current.

$e_a e_g$ plane in the line OM of which the equations are $i = 0$, $e_a + \nu e_g = 0$. A line through O perpendicular to this line and the axis of i, but not shown in the figure, is the axis of lumped voltage. Curves OA and OG are drawn in the figure to indicate the sections made on the surface by a plane through the e_a axis and a plane through the e_g axis. The curve OA is the $i_a e_a$ characteristic for zero grid voltage, for here the grid current is negligible and therefore $i = i_a$. The curve OG is the $i_g e_g$ characteristic at zero anode voltage; the curve OL is the lumped characteristic of the total current i. The lumped characteristic is the steepest curve of section and is the shape seen when the surface is viewed in profile.

157. On the left-hand side of the curve OG the anode voltage is negative and the anode current therefore zero. In consequence, this portion of the surface relates solely to the grid. On the right-hand side of the curve the grid voltage may be positive or negative; in the latter case the grid current is zero and therefore the surface relates to the anode. In making these statements it is assumed that the currents due to initial velocities are negligible in comparison with those utilised normally. It follows that sections made by planes parallel to the OA and OG planes in the

regions just delimited yield $i_a e_a$ or $i_a e_g$, or $i_g e_a$ or $i_g e_g$ characteristics, or at least portions of these characteristics. The remaining portions of the characteristics may be obtained when the remaining parts of the characteristic surface have been separated into its anode and grid components.

158. This separation can be accomplished by aid of the various characteristics discussed in preceding paragraphs. Fig. 286 shows the grid characteristic surface, its shape being conveyed by means of section lines made by planes parallel to the e_a and i axes; that is to say, the curves on the surface are $i_g e_a$ characteristics. The contour lines of such a surface are seen in Fig. 284.

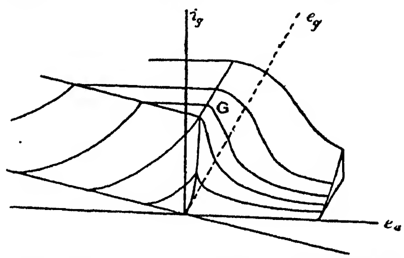


FIG. 286.—Characteristic Surface for the Grid Currents.

The anode characteristic surface may be drawn in the same way by aid of preceding paragraphs or by subtracting the surface of Fig. 286 from that of Fig. 285. It will be left to the imagination of the reader.

SOFT TUBES.

159. We have discussed the triode so far as if it were a perfectly evacuated tube and the current through it purely electronic. We must now glance at some very useful forms of tube in which there is an amount of gas sufficient to affect their properties. Such tubes are called "soft," while those highly evacuated are called "hard." The general effects of traces of gas are easily deduced from § 88, which deals with the phenomena in diodes. The steep rise of the current, after ionisation starts, now occurs in two circuits instead of one; in fact, the increase of current produced in a diode by the incursion of positive ions into the neighbourhood of the cathode is now divided between two cold electrodes, the proportion snatched by the grid being a function of the ratio of the voltages of the electrodes. The total current obeys the same rules as in a diode; for instance, it cannot exceed the emission corresponding to the temperature of the filament unless positive electrons reach the filament and deliver their charges to it in appreciable numbers—which is, as has been explained, a rare event.

160. Assuming in the first place that the grid voltage is below the ionising potential of the gas in the tube we may notice that the anode and grid voltages have very different effects in provoking ionisation. Consider the motion of an electron along a radius of a cylindrical triode from filament to cylinder, the radius being placed so that the electron passes midway between two turns of the grid, and let the grid be maintained at a moderate positive voltage, say 10 volts. Let the gas be helium and the anode voltage be a little above the ionising potential, say 22 volts. Then while the electron moves from the filament towards the grid it is accelerated by the fields due to both electrodes, but while it moves between grid and anode it is retarded by the grid's field and accelerated by the field of the anode. In a tube of ordinary dimensions the external retardation exerted by the grid practically undoes the acceleration produced inside the grid, whatever the voltage of the grid. Thus it is the anode voltage alone that raises the energy of the electron to the ionising value; hence, if in an experiment the grid voltage be fixed at less than the ionising potential and the anode voltage be raised gradually, ionisation by collision starts at the same anode voltage whatever the voltage of the grid. When, in helium for example, the anode reaches about 22 volts a few positive ions are formed near the anode, are repelled towards the cathode, assist the existing field to cancel the space-charge field, and therefore increase both the grid and anode currents. As the anode voltage is raised further the number of positive ions produced per second increases because the necessary electron velocity is attained further from the anode and therefore a greater volume of gas is subject to bombardment; in consequence the anode and grid currents increase together as the anode voltage is increased. This is illustrated in Fig. 287.

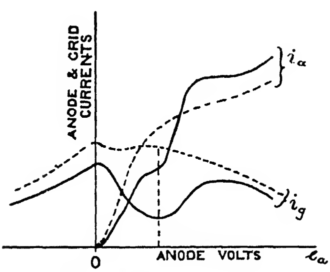


FIG. 287.—Characteristics at Low Anode Voltage.

161. Although, as just described, the grid voltage, if fixed below the ionising potential, does nothing to help the anode voltage in starting ionisation, yet it has great influence on the

amount of ionisation; for the field that liberates the electrons from the space-charge field is proportional to the lumped voltage, and therefore the higher the fixed grid voltage, within the limit stated, the more numerous are the flying electrons and, consequently, the more abundant the ionisation; that is to say, the steeper are the rising parts of the anode and grid gas lines. This is subject to the adequacy of the supply of electrons, for obviously, if the temperature of the cathode is so low that all the electrons are drawn from it before the anode reaches the ionising potential, neither grid nor anode current will experience any sudden rise.

162. In the above paragraph the grid was assumed to be at a voltage less than the ionising potential of the gas in use; let us now consider the case when the grid voltage is fixed slightly above the ionising potential, and let the anode voltage be again gradually raised from zero. Even at zero anode voltage positive ions will be present to assist the grid to neutralise the space-charge field and therefore a large grid current may be expected from the beginning. As the anode voltage increases, the rate of ionisation at the grid will not increase much, and therefore the anode current increases only slowly. But when the anode

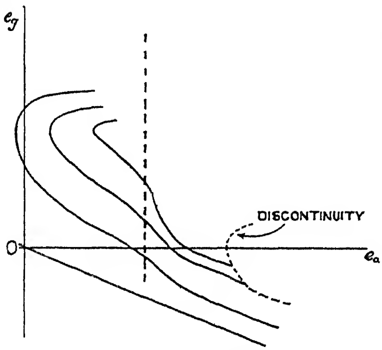


Fig. 288.—Anode Current Contours of Soft Tube.

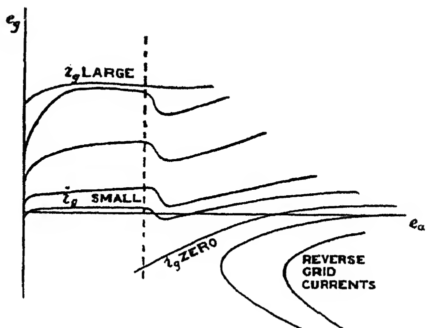


Fig. 289.—Grid Current Contours of Soft Tube.

voltage attains the ionising potential the space near the anode will become an additional source of positive ions and a rapid rise of both currents will, in general, occur. The trend of the phenomena is shown by the curves in Fig. 287. As before, the sum of the currents at any stage does not exceed the emission from the cathode.

163. Another mode of exhibiting the effects of gas in a triode is given in Figs. 288 and 289, which are sketched by aid of figures given by B. S. Gossling in a paper read before the Institution of Electrical Engineers in June, 1920. The curves are constant current lines, or $e_a e_g$ characteristics, that is to say, they are contour lines on the characteristic surface. They enable us to picture the distortion of the anode and grid characteristic surfaces by the presence of gas. Still another mode of representation is shown in Fig. 290, which relates to Round's C valve containing nitrogen at 0·31 mm., and in Fig. 291, which belongs to an R2 valve containing nitrogen at 0·1 mm, both of which are taken from the paper cited. These latter curves are perhaps the most familiar and most useful form of characteristic. In comparing them with the corresponding characteristics of hard tubes it is seen that the $i_a e_g$ curve is steeper when gas is present, and that when positive ions are present a substantial reverse current flows in the grid circuit when the grid emf is reversed. In Fig. 291 it is seen that the grid current marked 23·5 vanishes at about a quarter volt negative. The reversal of the grid current when the grid is made negative indicates that positive ions are being collected by the grid faster than electrons. As the grid potential is increased negatively the outward current from the grid increases to a low maximum and then decreases. The increase is

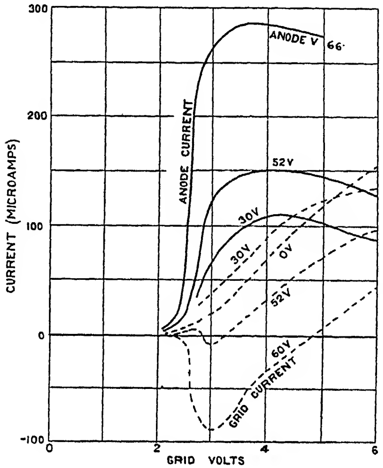


FIG. 290.—Characteristics of Round's C Valve (Gossling).

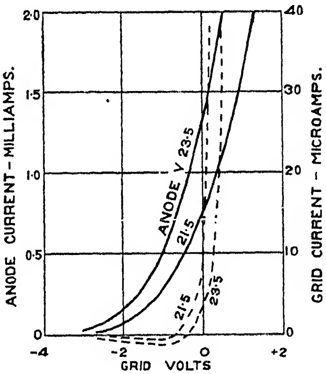


FIG. 291.—Anode Current (full line) and Grid Current (broken line) of Soft Tube (Gossling).

to be expected so long as positive ions continue to be formed by collision between molecules and high speed electrons in the space between grid and cylinder, and the decrease sets in because of the increasingly negative grid permitting continually fewer electrons to pass through it to carry on the ionising process. This falling curve expresses a noteworthy state of things—the state called "negative resistance," which is discussed in later pages.

164. The contrast with hard tubes offered by the existence of an outward current from the grid when this is negative relative to the filament—a phenomenon sometimes called "backlash"—is a good criterion of the presence of gas in a tube and shows, roughly, its amount. This current is greater at high anode potentials than at low for the reason that there is more ionisation by collision at high potentials, as explained in § 160. When the grid is positive, on the other hand, it collects no positive ions, and, just as in the case of hard tubes, it catches fewer electrons the greater the anode voltage. On both counts, therefore, the whole grid current curve is moved down the page by increasing the anode voltage. This is shown very clearly by the dotted curves marked 21·5 and 23·5 in Fig. 291. In consequence of this movement the point of zero current, which may be called the grid equilibrium voltage, shifts to the right when the anode voltage is raised, and conversely, shifts to the left when that voltage is lowered. By adjusting the working anode voltage, therefore, the equilibrium voltage of the grid can be made positive, negative or zero at will.

Adjustment of the Equilibrium Voltage.

165. A similar and more delicate adjustment can be accomplished by varying the filament voltage. Increase of this voltage has the effect of making the part of the grid confronting the positive end of the filament relatively more negative and therefore on the whole the grid becomes more negative relative to the filament and the ordinates of any grid curve are all diminished. On the other hand the raising of the temperature of the filament causes the grid to receive more electrons travelling under their initial velocities. And where the end effects are great the raising of the filament current may make active again

a short length formerly saturated and so augment the grid current. The balance in small reception tubes is usually in the direction due to initial velocities, that is to say, the grid curves are lifted slightly by the increase of filament voltage.

166. This definite possibility of adjustment is of great value in the use of the triode as a detector. As will be more fully explained later the heel of the $i_g e_g$ characteristic is usually used for rectification. Now the effect of the rectification of a train of oscillations at the grid is repeated in the anode circuit in the form of a transient change of current which is greater the steeper the $i_a e_g$ characteristic. From Fig. 291 we know that these latter characteristics move to the right as the anode voltage is diminished, which is the direction opposite to that taken by the grid curve and its heel. Therefore it is possible to bring the heel of the grid curve into association with the steepest part of the $i_a e_g$ curve. Further, in a well-designed and well-made tube, not only can the above condition be fulfilled, but also the equilibrium voltage of the free grid can be brought to zero voltage by adjustment of the filament current.

167. We may glance at another contrast between hard and soft tubes, which is of importance when the grid is used insulated, that is, connected to the insulated plate of a condenser. In a hard tube when the grid is used for rectification it collects negative electricity, which cannot escape. This accumulation of electricity will at least change the adjustments if it does not stop the anode current entirely. In the soft tube this changed condition can only be very transient, for the negative charge on the grid attracts positive ions and is thus neutralised. In order to enable a hard tube to be used for rectification in the manner described the grid condenser has to be short-circuited by a suitable discharging resistance, called a grid leak, which may be some hundreds of thousands of ohms; and this does what the soft tube does intrinsically. This property of the soft tube is, however, not always an advantage, for adjustments are sometimes attained in which the charging and discharging proceeds automatically at such a rate as to produce current changes of acoustic frequency in the respective circuits—to produce, in fact, in the connected telephones a howl which may range in pitch from a mere throb or rattle to a shrill shriek.

Glow Discharge.

168. When a sufficient voltage is applied to a tube containing more than a trace of gas the occurrence of ionisation by collision is made evident to the eye by the appearance of a blue glow. In these circumstances the behaviour of the tube is erratic for the reasons explained in discussing diodes, and, since bombardment of the filament raises its temperature and the emission, instability may ensue and the discharge pass over into an arc unless external ballast resistance is employed.

Some designs of soft tube when employed singly in the receipt of signals gave their best performances in the blue glow stage. Some of the early Round tubes were instances of this. Other soft tubes, such as de Forest's audions, worked best just below the blue glow stage. Another soft tube, White's pliotron, has in it a very minute trace of mercury vapour which is kept at practically constant pressure by the presence of mercury-silver amalgam in the tube. This tube is not so soft as the others mentioned, for it has been stated that 200 volts could be used on the anode without the slightest glow being produced.

General Equations of Triode Circuits.

169. The anode current i_a and the grid current i_g can each be measured when the anode voltage e_a and the grid voltage e_g have any assigned values. This may be expressed by writing.

$$i_a = f_a(e_a, e_g)$$
$$i_g = f_g(e_a, e_g).$$

We shall take e_a and e_g as independent variables. We then have

$$di_a = \frac{\partial i_a}{\partial e_a} de_a + \frac{\partial i_a}{\partial e_g} de_g$$

and

$$di_g = \frac{\partial i_g}{\partial e_a} de_a + \frac{\partial i_g}{\partial e_g} de_g$$

or

$$di_a = a_a de_a + a_g de_g$$
$$di_g = g_a de_a + g_g de_g$$

according to the notation already explained.

170. Now when the triode is connected into any circuit comprising inductance, resistance, capacitance and battery associated in any manner the result is to establish between the four variables of the triode at least one relation of the form

$$F(i_a, i_g, e_a, e_g) = 0,$$

whence $$\frac{\partial F}{\partial i_a}\, di_a + \frac{\partial F}{\partial i_g}\, di_g + \frac{\partial F}{\partial e_a}\, de_a + \frac{\partial F}{\partial e_g}\, de_g = 0.$$

This and the two differential equations above furnish equations for the solution of every problem concerning the triode and its circuits. In particular, since from the three equations we can eliminate two of the variables we may proceed to establish relations between any chosen pair of variables. For example, if we require to know how i_a changes in the complete circuit in consequence of a variation of e_g, we eliminate di_g and de_a from the equations. We obtain

$$\begin{vmatrix} 1 & 0 & -a_a \\ 0 & 1 & -g_a \\ \dfrac{\partial F}{\partial i_a} & \dfrac{\partial F}{\partial i_g} & \dfrac{\partial F}{\partial e_a} \end{vmatrix} di_a = \begin{vmatrix} 0 & -a_a & a_g \\ 1 & -g_a & g_g \\ \dfrac{\partial F}{\partial i_g} & \dfrac{\partial F}{\partial e_a} & -\dfrac{\partial F}{\partial e_g} \end{vmatrix} de_g$$

171. As another example, we may deduce an expression for the current amplification of the assemblage. It is obtained by eliminating de_a and de_g from the equations. We obtain

$$\frac{di_a}{di_g} = \begin{vmatrix} 1 & -a_a & -a_g \\ 0 & -g_a & -g_g \\ \dfrac{\partial F}{\partial i_a} & \dfrac{\partial F}{\partial e_a} & \dfrac{\partial F}{\partial e_g} \end{vmatrix} \div \begin{vmatrix} 0 & -a_g & -a_a \\ 1 & -g_g & -g_a \\ \dfrac{\partial F}{\partial i_g} & \dfrac{\partial F}{\partial e_g} & \dfrac{\partial F}{\partial e_a} \end{vmatrix}$$

Other Ways of Using Triodes.

172. In the preceding pages the triode tube has been supposed to possess a straight filament, surrounded by a net or a helix of wire forming a grid, which is again surrounded by a cylindrical tube of metal forming the anode. But it should be noticed that these relationships may be interchanged. For instance, when the grid is a helix with two external leads it may be made the hot cathode, and the straight axial filament may then be used as the grid was formerly, the cylinder still being the anode. This arrangement would have many of the properties of the form already discussed—for the positive charge on the axial conductor will, up to a more or less definite potential, clearly tend to neutralise the field of the space-charge formed round the glowing helix and so will liberate electrons for a journey to the anode. In view of this possible mode of using a tube, it is frequently advisable to give the name "control electrode" to that member used for neutralising the space-charge and so controlling the

current to the anode. Again, it would not be impossible to use the metal cylinder as the hot cathode, the axial conductor as the anode and the grid as control member; but there might be no advantages and there would be many disadvantages in such an arrangement.

The Dynatron.

173. Up to the present in discussing the triode it has been supposed that the potentials of the electrodes and other circumstances were such that the phenomenon of secondary emission was not provoked. The phenomenon has been referred to in § 10, but must here be touched upon further.

Let the grid of a triode be raised to a higher voltage than the cylinder, both voltages being fairly high. Then it is evident that relatively high speeds will be attained by the electrons between filament and grid, that is, in the space where formerly very low speeds were the rule. The bulk of the electrons will, as before, shoot past the grid and will strike the cylinder with high velocity. This results in the production of other electrons, apparently knocked out of the atoms of the metal of the cylinder by the impact of electrons from the filament. These are sometimes called secondary electrons in contradistinction to the original electrons, which are called primary; but the new electrons were called δ rays by their discoverer, J. J. Thomson. No doubt many of these secondary electrons were produced during the already described modes of operating the tube, but they were unobserved because attracted back into the cylinder after a short flight. The δ rays come into notice with the new dispositions merely because there is now a neighbouring electrode at higher potential, namely, the grid, which attracts them to itself and so causes them to be missed from the cylinder. A. W. Hull has shown that high speed primary electrons may produce by impact on metal so many electrons that a positive grid near the cylinder collects twenty secondary electrons on the average for each primary one incident on the plate.

174. The δ rays, if collected by the very positive grid, are, as has been said, missed from the cylinder. The constant voltage battery connected to the cylinder must supply the deficiency. Thus there is a net negative current from outside towards the

cylinder equal to the secondary electrons emitted minus the primary electrons received. It is to be expected, and it is found on trial, within limits, to be the fact, that if the potential of the cylinder be raised, though still kept below that of the grid, the speeds of impact of the electrons and therefore the rate of emission of δ rays will be increased. But the rate of production of primary electrons by the filament is quite unaffected by the voltage of the cylinder. Therefore the net electron current from the cylinder to the vacuum is increased. Thus, raising the voltage between cylinder and filament produces an increase of current just as in a conductor, but in the direction opposite to the voltage. The tube is therefore said to possess negative resistance. The tube and its battery constitute a device which supplies to the circuit energy varying in proportion to the square of the current traversing it.

Negative Resistance.

175. This term is applied to any piece of apparatus or machinery that acts as a source of emf in such a way that its emf, or the variable part of it, is proportional to the current through the tube and is in the same direction. The multiplier in this proportion is the " negative resistance " coefficient. Thus if e represents the variable part of the emf, i the current, and r the resistance coefficient we have

$$e = -ri.$$

This way of regarding the device has certain mathematical advantages, but it is usually simpler physically to look at the device as it really is, namely, a source of emf of strength e determined by the current traversing it, being in fact r times that current.

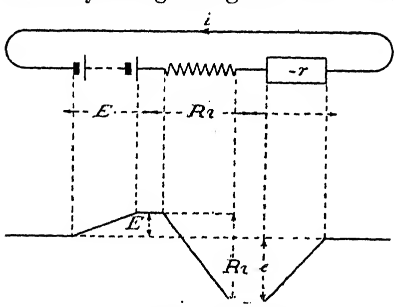


FIG. 292.—Diagram of Voltages in Circuit containing Negative Resistance.

The properties of such an apparatus may be expressed graphically by drawing a diagram of voltages round a circuit comprising a voltaic battery, a resistance coil, and a negative resistance. This is done in Fig. 292.

The Dynatron.

176. The delta rays just described have been made by A. W. Hull the basis of a type of triode called the dynatron, which

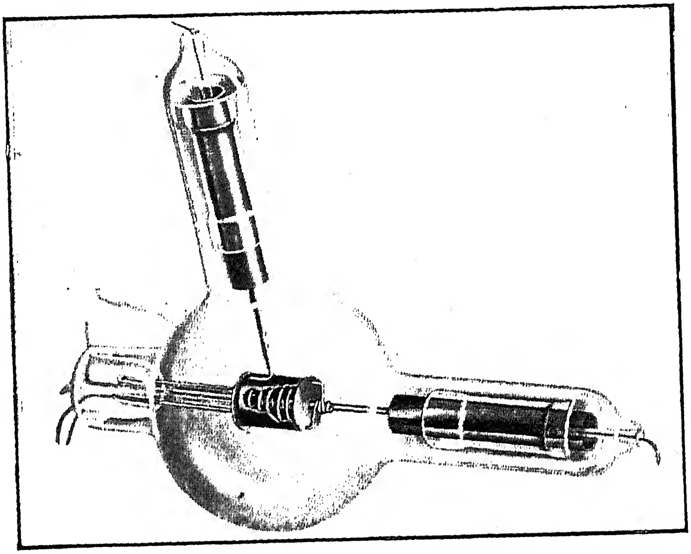


FIG. 293.—The Dynatron (from the Proceedings of the Institute of Radio Engineers).

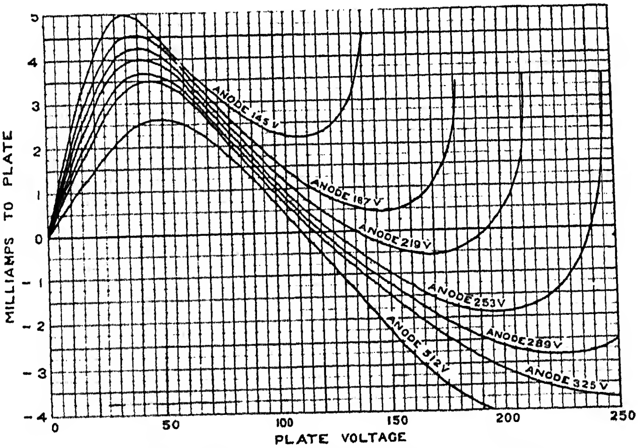


FIG. 294.—Characteristics of a Dynatron (Hull).

has been developed in the General Electric Laboratory. The construction is seen in Fig. 293. The cylinder is shown cut open

to expose a helix of stout wire between the cylinder and the glowing filament. Instead of a helix of stout wire this member

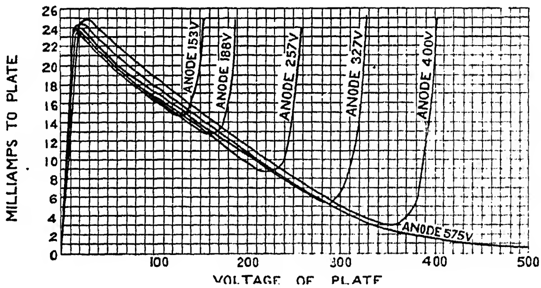


FIG. 295.—Characteristics of a Dynatron (Hull).

may be a perforated tube or a fine tungsten wire cage. The hot cathode is a close-wound helix of tungsten wire. Since the most essential function of the cylinder is connected with its bombardment by electrons, we may speak of it as the target. In some of the figures the cylindrical target is called the plate. The helix or perforated plate or wire cage is the electrode maintained at highest potential and is therefore called the anode. Figs. 294 and 295 are taken from a paper by Hull, and show the characteristics of two tubes differing only in their geometrical forms. The abscissæ are the voltages applied to the target, and the ordinates are the current flowing into the target from outside the tube. The anode, that is to say, the helix, is maintained

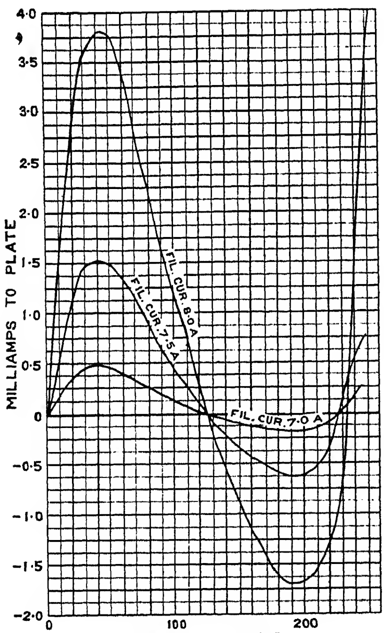


FIG. 296.—Effect of varying Filament Current of Dynatron (Hull).

at a constant potential throughout each curve, the value of the potential being marked on the diagram. The effect of varying

the anode voltage is decidedly different in the two tubes, but in each case the range of the negative gradient is wider when the voltage is greater. In Fig. 294 the straight parts of the curves move to the left with increasing anode voltage. In Fig. 295 they move to the right, which indicates that it would be possible to design a tube in which the straight parts of negative gradient would coincide at every anode voltage. Within certain values the negative resistance is fairly constant, being about 18,000 ohms in Fig. 294 and 12,500 ohms in Fig. 295. In Fig. 296 the effect of varying filament current is shown. The range of the negative gradient is not much changed by change of filament current but its magnitude changes rapidly, being about 200,000 ohms for the smallest filament current, 60,000 for the medium and 25,000 for the largest filament current. The various uses of the dynatron will be discussed in the proper sections.

Pliodynatron.

177. Instead of varying the temperature to produce changes

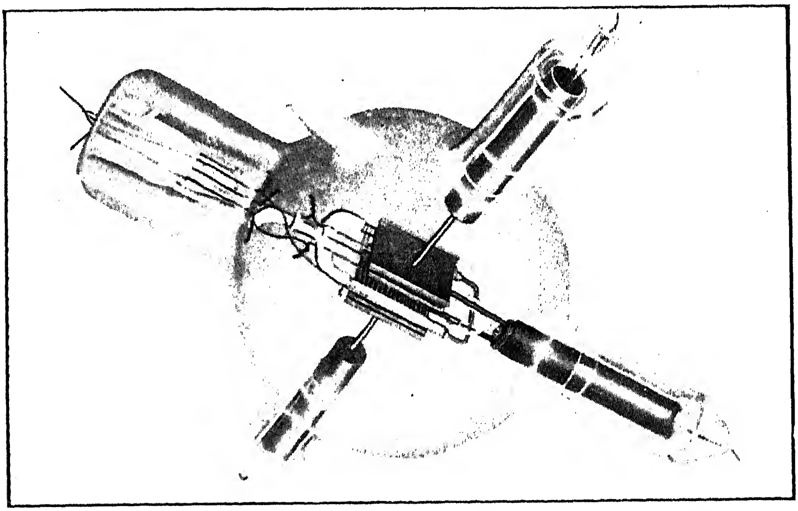


FIG. 297.—Pliodynatron (from the Proceedings of the Institute of Radio Engineers).

in the number of electrons leaving the filament, such changes can be produced by means of a control member introduced into the tube in such a manner as to perform the same functions as

the grid in the pliotron. We thus obtain a four-electrode tube, that is to say, a tetrode. The instrument developed by the General Electric is called the pliodynatron. A photograph taken from the *Proceedings of the Institute of Radio Engineers* is reproduced in Fig. 297. There is a perforated plate between the fine wire grid and each of the outer plates, and the two together form the high potential member or anode. In some tetrodes the control electrode is not the grid surrounding the filament, but is a metal rod inside a helical filament, and then the rod behaves in the manner described in § 172.

Assuming that the tube is operated on the straight part of the ordinary triode characteristic, the electron current leaving the filament will vary with the difference of potential between grid and filament; while the negative resistance of the dynatron will vary inversely with the total number of electrons leaving the filament. Consequently the negative resistance between target and filament is inversely proportional to the grid potential.

The use of the tube for magnifying will be described in another section.

General Note.

178. In the treatment of the ionic tube developed in this chapter the author has been greatly assisted by the various publications of Irving Langmuir in the *Physical Review*, the *Annalen der Physik*, and especially by those published in the *General Electric Review* of June and July, 1920. These papers deal with the theory of the space-charge and with the effects of the admission of gas upon the back emf of the space-charge. Papers by S. Dushman have been consulted for information about the kinetic theory of gases and about the emission from filaments. Readers interested in the history of thermionic tubes, especially with regard to their development in this country, should refer to a paper by B. S. Gossling, read before the Institution of Electrical Engineers, June, 1920.

On the other hand, for the theory of the control electrode, the grid current and the characteristic surface, and for the discussion on the effects of non-uniformity of filament temperature and potential, the author is mainly responsible. It should be mentioned that much of the fundamental mathematical treatment of the triode as a whole, without regard to internal actions of the tube, will be found in their essentials in M. Latour's paper in the *Electrician* of December 1st, 1916.

MEASUREMENT OF THE PARAMETERS OF TRIODES.

179. When a tube is about to be put into use we require to know one or two commonplace facts concerning it, such as the appropriate filament current and voltage, which are easily found out or can be supplied by the maker. But there are other properties of the tube—such as the voltage ratio under conditions of normal use—which usually have to be specially measured and are essential for the formation of an accurate forecast of the performance of the tube in particular applications, for instance, in an amplifier or in a receiving set. A large number of parameters have been suggested for expressing numerically the various properties of triodes, but not all of them are equally important nor are all of them independent—that is to say, some of them can be deduced from others. By careful selection it may in time become possible to convey by quite a small number of parameters practically all the information requisite in practice. We shall in fact omit from the present section a number of suggested parameters or "constants" which depend jointly on properties of a tube and of the circuits it happens to be placed in, and we shall confine our attention to parameters belonging purely to the tube. Mixed functions of the tube and its circuits can usually be deduced from the data of the circuits and the true parameters of the tube itself.

180. The principal parameters of a triode are :—

(1) The co-ordinates of the mid-point of the straight part of the lumped characteristic.

(2) The voltage factor ν.

(3) The differential conductance $\partial i_a/\partial e_a$ or a_a.

(4) The differential conductance $\partial i_a/\partial e_g$ or a_g.

If these are all supposed to relate to the regular parts of the lumped characteristic and not to extreme and unusual adjustments the last-named parameter is not independent of the preceding two since $a_g = \nu a_a$. Other data of importance are :—

(5) The differential conductance $\partial i_g/\partial e_a$ or g_a.

(6) The differential conductance $\partial i_g/\partial e_g$ or g_g.

(7) The derivatives $\partial i_a/\partial e_f$ or a_f and $\partial i_g/\partial e_f$ or g_f.

In addition, data concerning the bends of the characteristic curve are of value in connection with the use of the tube as a detector.

Among other useful information about a triode may be

included the grid potential at which appreciable grid current begins to flow, and the potential to which the grid tends when insulated, that is the equilibrium voltage of § 164, the anode voltage being normal. The anode current at this setting may also be useful.

All these parameters belong wholly to the tube, that is to say they are not affected in value by the presence of the apparatus with which the tube is used and are, so far as is investigated independent of the frequency. From a knowledge of these magnitudes the main aspects of the behaviour of the triode when connected to any kind of apparatus can be calculated. All of them can be deduced from the characteristic curves of a tube but as the drawing of a number of complete curves is laborious we proceed to indicate various methods of determining directly the most important parameters.

181. The quantities ν, a_a and a_g, as well as some of the others are either mere ratios or are of the nature of conductances. Their measurement can therefore be reduced to the comparison of resistances, which can by the aid of resistance boxes be accomplished conveniently and accurately. But several of the measurements, notably those of parameters defined by differential coefficients, are concerned with small changes in relatively large currents or voltages, and it is evident that if all of a large current is passed through a galvanometer the sensitiveness for small changes must be low. When accuracy is desired we therefore adapt some of the general methods of the physical laboratory to our purpose, methods such as the various resistance bridges, the compensation method, or methods using the differential galvanometer, in all of which the galvanometer can be maintained in its most sensitive adjustment and be kept near its zero although large currents are running in the things being measured. These and some other methods which have proved useful in the author's laboratory, as well as methods to be found in various published papers by different authors, will be described in the succeeding paragraphs. Most of them are essentially direct current methods or can be used with direct current.

D.C. and A.C. Methods.

182. Direct current methods have certain advantages over methods using rapidly alternating currents and a telephone

receiver, though they are not so quick and convenient. An alternating current gives rise to alternating electric and magnetic forces round the various portions of its circuit, and these forces excite alternating emfs in neighbouring conductors and also in the different parts of its own circuit; in other words, the accidental electric coupling and magnetic coupling between portions of the same and other circuits produce emfs, and these add subtly to the emfs applied for the purpose of the measurement. The effects are greater the higher the frequency of the alternations, and are usually too easily evident at ordinary acoustic frequencies. If the resultant of the unwanted emfs happens to be in phase with the applied emf it is possible to obtain silence in the telephone, but the balance adjustment so obtained will usually be farther wrong the stronger the induced stray effects. If, on the other hand, the resultant of the stray emfs is not in phase with the applied emf it is impossible to obtain a perfect balance.

In order to minimise these alternating current disturbances it is well to run as much as possible of the connecting wiring through metal pipes or to wind the insulated wires with lead or tin foil, and to connect all this metal covering to earth. It is necessary also to place coils in such relative positions that they have the least possible coupling with other coils, and to place condensers and other conductors of large surface well away from one another.

Co-ordinates of the Mid-point.

183. The most obvious and direct method for determining the lumped voltage and the corresponding current at the mid-point of the characteristic curve is to apply known grid voltage and anode voltage, to increase the latter until the saturation current is reached, to divide the magnitude of this by two, and then reduce either grid or anode voltage until the current becomes half of the saturation value. But the last step demands the use of calibrated sources of voltage, and as these are not always at hand the following less direct method is recommended.

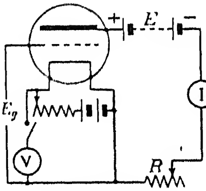


Fig. 298.

The connections are shown in Fig. 298. The galvanometer I

ought to be of low resistance compared with that offered to the current by the internal path between anode and cathode. The filament emf or current must first be set at the desired stationary value, by aid of the voltmeter V or otherwise, and the variable resistance R made zero. Then make the grid voltage zero or other desired value and increase the anode voltage E by large steps till the current, as registered by the galvanometer I, enters the saturation stage. Observe these values of anode voltage and current and then add resistance R to the anode circuit until the galvanometer reading is reduced to one-half the saturation value, say, i_1. Then the required co-ordinates are i_1 and $e_1 = E - Ri_1$. It should be noticed that this simple experiment gives information of great utility in connection with trials for amplifiers, for the voltage e_1 is that lumped voltage which must be applied to the tube in order to bring the steepest part of the characteristic into use; and if the voltage on the grid is to be kept zero, then e_1 is the voltage which must be applied to the anode. As an example, in a certain tube with a filament voltage of 5·2 V an anode voltage of 300 secured a saturation current of 10 mA. By adding $R = 28000\ \Omega$ in series the current was reduced to 5 mA, the grid being connected all the time to the negative side of the filament. Therefore $e_1 = 300 - 28 \times 5 = 160$ V. The result is, it should be noticed, a lumped voltage; it may be made up by any anode voltage and grid voltage that satisfy the equation

$$e_a + \nu e_g = 160.$$

For example if $\nu = 10$

we may have

$e_a = 100$	and $e_g = 6$
130	3
160	0
190	-3
220	-6

Any of these settings will give practically the steepest part of the characteristic.

In many tubes the saturation current is ill-defined and, indeed, the current may increase indefinitely as the anode voltage is raised. But for the present purpose it is sufficient if the anode voltage be taken to a value at which the addition or subtraction of, say, 20 volts, makes a change in the current which is small relative to that occurring at medium voltage.

184. Alternatively we may search for the steepest place on the characteristic by alternately adding and subtracting a small emf either in the grid or in the anode circuit, and reading the increase or the decrease of anode current. The circuit shown in Fig. 299 has the reversible battery in the anode circuit. Suppose the anode, grid and filament voltages to have any assigned values, then if on performing the experiment the increase of anode current is greater than the decrease, we conclude that the existing adjustment corresponds to a point on the heel of the characteristic curve; if the increase is less than the decrease the point is on the knee, and if the increase is equal to the decrease the point is on the steepest portion.

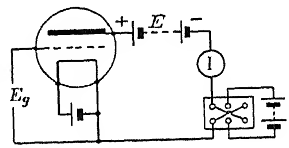


FIG. 299.

Measurement of Voltage Factor and Resistances of Triodes.

185. The possible methods of measurement of the chief triode parameters are very numerous and a large number of methods have been published; the methods to be described in the next six paragraphs possess the merit of being very quick and of requiring only very simple apparatus. They are, of course, not so accurate as the more elaborate methods described in the succeeding paragraphs.

Voltage Factor ν.

186. When a potential difference e_g is applied between the grid and filament the internal back emf due to the space-charge is partially annulled and, therefore, some of the voltage of the main anode battery is set free in the anode circuit. This liberated emf is, we have seen, of amount νe_g, where ν is the voltage factor, and it will produce an additional anode current unless the voltage applied in the anode circuit be correspondingly reduced in some way. In fact, in order to keep the anode current constant the lumped voltage on the tube must be kept constant, the lumped voltage being the sum of the applied anode voltage and the transferred voltage νe_g.

187. In the method indicated by the circuit diagram of Fig. 300 the resistance R has a constant value of about 10 ohms and r is a slide wire; the instrument I is a milliammeter and the

batteries E, E_g have the steady voltages at which the voltage factor has to be measured. The operation consists in adjusting the sliding contact on the slide wire until opening and closing the key K causes no change in the reading of the milliammeter. The closing of the key introduces a positive voltage ri' into the grid circuit and (very approximately) a negative voltage Ri' into the anode circuit, i' being the current produced in the auxiliary circuit by its battery. Since the anode current is unaltered we have

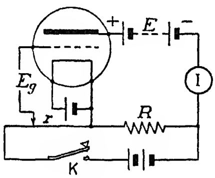


FIG. 300.

$$\nu r i' = R i'$$

or

$$\nu = R/r.$$

This method was given by E. V. Appleton in the *Wireless World* of November, 1918. The accuracy of the method may, if desired, be increased by using a differential galvanometer at I with its second winding carrying a compensating current. The slide wire may be graduated to read ν directly. The auxiliary current should be large compared with the triode currents.

188. The method indicated in Fig. 301 requires even less apparatus than the last method; the portion marked R is now a resistance box, preferably of dial pattern. The operations consist, first, in giving R any convenient value, say R_1, and observing i_a while the switch is on the point marked 1, and, second, putting the switch on the point 2 while adjusting the resistance to bring the anode current to its previous value. Let this second value of the resistance be R_2 and the anode current be i_a. Then on switching over from 1 to 2 the grid voltage rises by the amount $R_1 i_a$, and on adjusting the resistance to R_2 the voltage drop in the resistance box is increased by $(R_2 - R_1)i_a$. For constant lumped voltage we must therefore have

$$\nu R_1 i_a = (R_2 - R_1) i_a$$

or

$$\nu = \frac{R_2 - R_1}{R_1}.$$

For the quick measurement of ordinary reception triodes it is very convenient to make $R_1 = 1{,}000$, then the final reading of the resistance box minus 1,000 gives ν by the movement of a decimal point.

Numerical Example.—Filament current = 0·70 A. Anode volts = 63·0. Grid volts = 0·0.

R_1	1,000	900	800	700	600	500	400	300	
R_2	7,959	7,210	6,350	5,530	4,790	4,000	3,250	2,400	
ν	6·95	7·0	6·88	6·86	6·84	6·9	7·0	6·9	Average 6·9

By parting R into two boxes and providing an extra switch, as suggested at the side of Fig. 301, the subtraction can be avoided. The box R_1 may consist merely of a thousand-ohm coil and $R_2 - R_1$ is then read on the dial box.

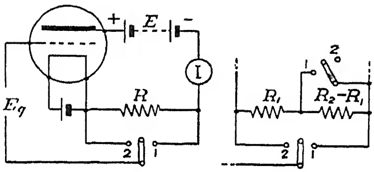


Fig. 301.

Anode Conductance a_a or its Reciprocal r_a.

189. The simple method to be given is almost sufficiently explained by the diagram of Fig. 302, where R is the only variable for the purpose of the measurement. Let S be the resistance of the instrument and battery and r_a the internal or differential resistance of the tube. First make $R = 0$ by closing the switch and observe the reading of the instrument, say i_1; then give to R a convenient value and read the changed value of the current, say i_2. Since the applied voltages are constant the currents are inversely proportional to the total resistance in circuit, that is

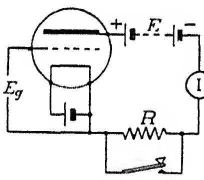


Fig. 302.

$$\frac{i_1}{i_2} = \frac{r_a + R + S}{r_a + S},$$

or

$$r_a = \frac{Ri_2}{i_1 - i_2} - S.$$

As a rule S is negligible compared with r_a and then

$$r_a \doteqdot \frac{Ri_2}{i_1 - i_2}.$$

A dial resistance box is best at R and it is convenient to proceed as follows: With $R = 0$ read i_1 (an arbitrary scale will serve), and divide it by 11; then adjust R until the reading is $i_2 = 10i_1/11$. This value of R multiplied by 10 is r_a.

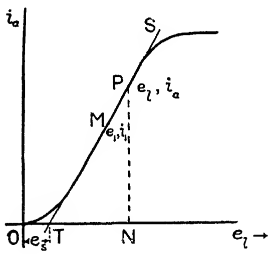


FIG. 303.

The most direct way of measuring the anode conductance is that suggested by Fig. 303, the apparatus being the same as that used for finding the co-ordinates of the mid-point. For if the resistance R in the anode circuit be given the value R' and the current be found to be i_a the lumped voltage becomes $e_l = E - R'i_a$ (if we assume $E_g = 0$), and therefore

$$e_l - e_1 = Ri_1 - R'i_a$$

and
$$a_a = \frac{Ri_1 - R'i_a}{i_a - i_1}.$$

The figure also shows that however a_a be measured the magnitude of the space charge back emf is

$$e_s = e_1 - i_1/a_a.$$

Measurement of a_g.

190. The method described by the diagram of Fig. 304 was given by E. V. Appleton in the *Wireless World* of November, 1918. When the key K is open the galvanometer is traversed by the anode current, and when K is closed the galvanometer is traversed in opposite directions by the changed anode current and the new auxiliary current through the resistance r. The change in the anode current is due to the P.D. created in r by the auxiliary current. Let this resistance be adjusted until the galvanometer reading is the same whether the key is up or down; then if i represents the auxiliary current and also the change of anode current to which it must

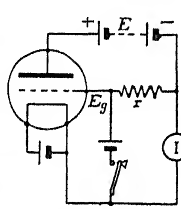


FIG. 304.

al, we have, when the resistance of the galvanometer e neglected,

$$a_g = \frac{\text{change of anode current}}{\text{change of grid voltage}} = \frac{i}{ri}$$
$$= 1/r.$$

[ethods.

J. M. Miller has described a method in which the principal its are measured by means of currents of audio frequency. cuit is given in Fig. 305.

neasurement of the ; ratio of the tube is l while the key K is y adjusting the ratio R_2 to give zero sound telephone. Clearly adjustment the emf to the grid by R_1 is

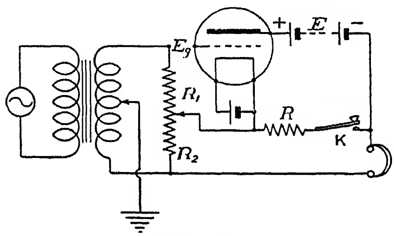


FIG. 305.

isated by that applied directly to the plate circuit by 1 therefore

$$\nu = \frac{R_2}{R_1}.$$

ose the key K, altering the ratio R_2/R_1 and varying R to silence again; then if we assume that R is so much larger 2 that the current i in the potential divider is unchanged, e from the circuit R K R_2

$$R_2 i = R i_a,$$

the anode, filament and R circuit (since there is no alter-current through the telephone)

$$\nu R_1 i = (R + 1/a_a) i_a.$$

and i_a are small audio currents. From these two equations ain

$$a_a = \frac{R_2}{R(\nu R_1 - R_2)}.$$

In the use of this method Miller recommends that the ce R should be a non-inductive dial box giving up to Ω, and that the potential divider $R_1 R_2$ should be a slide , for instance, 7 Ω resistance. Besides the earth connec-wn in the figure a variable earth connection on a rheostat,

connected in parallel with the slide wire, helps by sharpening the minimum. The current used in the potential divider may be less than 50 mA. Change of frequency does not affect the measured values of the tube constants. The telephone should be of rather low resistance and a voltmeter of high resistance should be used for determining the actual steady voltage on the anode. Some curves given with the description of the method in the *Proceedings of the Institute of Radio Engineers* of June, 1918, show that the voltage ratio ν, in a certain tube, varied only 2 per cent. when the anode battery voltage was changed from 50 V to 120 V, being 14·5 at 100 V. In the same tube at the same filament temperature the value of a_a followed approximately the straight line rule

$$a_a = 0{\cdot}29\,(E_a - 26)\,.\,10^{-6}\text{ mho}$$

between the steady anode voltages 50 and 110.

Compensation Methods.

193. It has been explained in preceding pages that the thermionic tube behaves as if it contained within it both resistance and back electromotive force; the methods of measurement known in the physics laboratory as compensation methods are therefore particularly appropriate for the accurate measurement of the parameters. In these methods, which are null methods, the normal current from the ionic tube may be regarded as passing through the galvanometer while an equal constant current passes through it in the opposite direction from an auxiliary source of current; the galvanometer may therefore be used in a sensitive adjustment, and very slight changes from the normal current may be detected or measured. When applied to the measurement of the principal parameters of a triode these methods are sensitive enough to appreciate the effects of very slight changes in filament and anode voltages. Conversely, it is very necessary that these voltages should be maintained very invariable when the parameters of the tube are being measured. The methods are applicable at any grid or anode voltages if the grid current is small. The connections of the apparatus for measuring a_a, a_g and ν are shown in Fig. 306. The anode battery is indicated by E, the grid potential divider by E_g, the galvanometer and shunt by I and S, and variable resistances by R, R_1, R_2, R_3. The compensation circuit contains the battery e, which is of smaller

voltage than E. The six-points switch when rocked to the left-hand side connects each of the middle points to one of the end points marked a and is then in position for the measurement of a_a; when rocked to the right it is placed for measurement of a_g. A key K is combined with the switch and is used in both the above measurements, but is opened for the measurement of the voltage factor. A telephone and interrupter may be substituted for the galvanometer in all the measurements, but then the precautions described in § 182 should be taken. Fuller details of the circuits will be found in the *Proceedings of the Physical Society*, 32, pp. 92—104, February 15th, 1920.

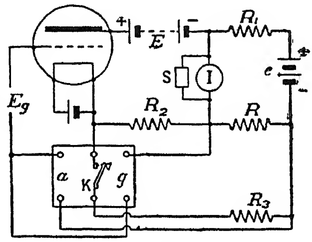


FIG. 306.

Determination of a_a.

194. Rock the switch to the left so that the grid is permanently connected to the filament throughout the experiment, set the various emfs at the desired values, and then proceed as follows: First, the galvanometer being shunted and the key K being closed, the resistance R_2 is given any value, R a small value (between 10 and 100 ohms in the case of a reception tube), and R_1 a larger value (about 1,000 ohms), and the latter is adjusted until the deflection of the galvanometer is zero. Then the key K is opened and R adjusted until the deflection is again zero. The shunt is now weakened, the key closed, and R_1 more accurately adjusted than before to obtain zero deflection; and then the key is opened and R improved in adjustment. The process is repeated to any required degree of accuracy.

195. Zero deflection of the galvanometer implies that the P.D. between its terminals is zero. Therefore, when the key is open the voltage applied in the anode circuit is $E - R_2 i_a$; similarly, when the key is closed the voltage is E. Here i_a represents the anode current when the key is open, and is equal to the current in the compensation circuit. Let e represent the voltage of the battery in this circuit, then

$$i_a = \frac{e}{R_1 + R}.$$

The current when the key is closed is e/R_1. Hence, on closing the

key the anode voltage increases by the amount $R_2 i_a$, and the anode current increases by

$$\frac{eR}{R_1(R_1 + R)} \text{ or } \frac{Ri_a}{R_1}.$$

Since the grid voltage is constant throughout the experiment the ratio of the increase of anode current to the increase of anode voltage gives

$$a_a = \frac{R}{R_1 R_2}.$$

Determination of a_g.

196. Rock the switch to the right. The grid is then connected to the filament direct when the key K is closed and through R_3 when the key is open. After the values of the grid, anode and filament voltages have been settled the measurement is performed as follows: The galvanometer being shunted and K open the resistance R is given a small value, and R_1 a larger value, which is adjusted till the deflection of the galvanometer is zero. The key K is now depressed and R_3 adjusted by trial until the deflection of the galvanometer is again zero. The shunt is then weakened, the key opened, and R_1 adjusted more accurately than before. The key is depressed and R_3 again adjusted. The process is repeated to any required degree of accuracy.

197. When the key is down the current in the compensation circuit and, therefore, that entering the anode is determined by R_1 in series with R and R_3 in parallel; when the key is up this current is determined by R and R_1 in series. In the former case the current being equal to e divided by the resistance is

$$\frac{e(R_3 + R)}{R_1 R_3 + R_1 R + R_3 R}$$

and in the latter case is

$$\frac{e}{R_1 + R}.$$

Hence the increase of current when K is depressed, being the difference of the last two expressions, is

$$\frac{R^2 e}{(R_1 + R)(R_1 R_3 + R_1 R + R_3 R)}.$$

Again, when the key is down the voltage on the grid is that of

the grid battery ; when K is up the grid voltage becomes smaller by the amount of the potential drop along R, namely,

$$\frac{Re}{R_1 + R}.$$

Moreover, the anode voltage is the same whether the key is up or down, since zero deflection of the galvanometer implies that the P.D. between its terminals is zero and, therefore, that E is the only voltage in the anode circuit. Hence, when the key goes from the open to the closed position the anode current and the grid voltage both rise, while the anode voltage remains constant, and therefore the ratio di_a/de_g is

$$\begin{aligned} a_g &= \frac{R}{R_1R_3 + R_1R + R_3R} \\ &= \frac{1}{R_1 + R_3 + (R_1R_3/R)} \end{aligned}$$

Determination of ν.

198. First, the compensation circuit is adjusted by variations of R_1, the switch being in the position marked g ; then the switch is rocked over and R_2 is adjusted till the deflection is again zero.

When the switch is in the position a the anode voltage is $E - R_2i_a$, and the grid voltage is that applied by the potential divider E_g. When the switch is moved to position g the anode current is the same as before because equal to the compensation current, which is unaltered. Moreover, in position g the anode voltage is E and the grid voltage is Ri_a below that applied by E_g. Thus the fall of the grid voltage Ri_a neutralises the rise R_2i_a of the anode voltage, and therefore

$$\nu Ri_a = R_2 i_a$$

or

$$\nu = R_2/R.$$

When the grid current is extremely great, relatively speaking, an error arises in the above measurement because current is drawn from the compensation battery to the grid. This is, however, a rare event. In such cases it is advisable to make R_3 zero during this measurement.

INDEX

THE WHITEFRIARS PRESS, LTD., PRINTERS, LONDON AND TONBRIDGE.